

GenCore version 5.1.6
Copyright (c) 1993 - 2004 Compugen Ltd.

OM protein - protein search, using sw model

Run on: February 19, 2004, 14:22:29 ; Search time 26.3838 Seconds
(without alignments)
3095.051 Million cell updates/sec

Title: US-10-042-431-49

Perfect score: 2076

Sequence: 1 VHMPTKAVDPFAFNISEII.....IHLMOQBETNLSQRCRAVL 390

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 801455 seqs, 209382283 residues

Total number of hits satisfying chosen parameters: 801455

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : Published Applications AA:
1: /cgn2_6/ptodata/1/pubaa/US07_PUBCOMB.pep:*
2: /cgn2_6/ptodata/1/pubaa/FCI_NEW_PUB.pep:*
3: /cgn2_6/ptodata/1/pubaa/US06_NEW_PUB.pep:*
4: /cgn2_6/ptodata/1/pubaa/US06_PUBCOMB.pep:*
5: /cgn2_6/ptodata/1/pubaa/US07_NEW_PUB.pep:*
6: /cgn2_6/ptodata/1/pubaa/PCTUS_PUBCOMB.pep:*
7: /cgn2_6/ptodata/1/pubaa/US08_NEW_PUB.pep:*
8: /cgn2_6/ptodata/1/pubaa/US08_PUBCOMB.pep:*
9: /cgn2_6/ptodata/1/pubaa/US09A_PUBCOMB.pep:*
10: /cgn2_6/ptodata/1/pubaa/US09B_PUBCOMB.pep:*
11: /cgn2_6/ptodata/1/pubaa/US09C_PUBCOMB.pep:*
12: /cgn2_6/ptodata/1/pubaa/US09_NEW_PUB.pep:*
13: /cgn2_6/ptodata/1/pubaa/US10A_PUBCOMB.pep:*
14: /cgn2_6/ptodata/1/pubaa/US10B_PUBCOMB.pep:*
15: /cgn2_6/ptodata/1/pubaa/US10C_PUBCOMB.pep:*
16: /cgn2_6/ptodata/1/pubaa/US10_NEW_PUB.pep:*
17: /cgn2_6/ptodata/1/pubaa/US60_NEW_PUB.pep:*
18: /cgn2_6/ptodata/1/pubaa/US60_PUBCOMB.pep:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	2076	100.0	390	11	US-09-759-130B-419
2	2076	100.0	390	14	US-10-042-431-49
3	2076	100.0	409	10	US-09-835-996A-21
4	2076	100.0	423	11	US-09-759-130B-417
5	2076	100.0	423	14	US-10-042-431-47
6	1991	95.9	374	10	US-09-811-825-4
7	1991	95.9	398	11	US-09-811-825-2
8	1289	62.1	399	11	US-09-759-130B-411
9	1289	62.1	399	12	US-10-051-874-92
10	1289	62.1	399	12	US-10-051-874-93
11	1289	62.1	399	12	US-10-051-874-94
12	1289	62.1	399	14	US-10-042-431-41
13	1282	61.8	392	14	US-10-003-302-4
14	1165	56.1	221	11	US-09-759-130B-420
15	1165	56.1	221	14	US-10-042-431-50

16	1161	55.9	395	12	US-10-015-115-98
17	1141.5	55.0	398	12	US-10-015-115-96
18	1141	55.0	357	12	US-10-108-260A-3705
19	1129	54.4	399	12	US-10-221-097-35
20	1127	54.3	378	12	US-10-420-564-2
21	1126	54.2	397	12	US-10-312-088-29
22	1109	53.4	371	12	US-10-015-115-97
23	1109	53.4	398	11	US-09-759-130B-445
24	1109	53.4	398	12	US-10-015-115-94
25	1109	53.4	398	12	US-10-015-115-95
26	1109	53.4	398	14	US-10-042-431-75
27	1082	52.1	398	14	US-10-056-744B-2
28	1024	49.3	401	12	US-10-015-115-32
29	1023	49.3	395	14	US-10-003-302-2
30	998	48.1	373	12	US-10-051-874-28
31	995	47.9	373	12	US-10-051-874-30
32	989	46.7	365	12	US-10-312-088-28
33	969	46.7	395	12	US-10-051-874-90
34	929	44.7	349	12	US-10-051-874-32
35	890.5	42.9	390	12	US-10-015-115-30
36	835	40.2	351	12	US-10-051-874-91
37	783	37.7	144	11	US-09-759-130B-422
38	783	37.7	144	14	US-10-042-431-52
39	400	19.3	127	14	US-10-056-744B-5
40	319	15.4	567	12	US-10-032-585-7720
41	296	14.3	476	15	US-10-128-714-3075
42	296	14.3	641	15	US-10-128-714-3075
43	294	14.2	431	12	US-10-304-038-6
44	209	10.1	86	12	US-09-864-408A-8008
45	128	6.2	25	11	US-09-759-130B-421

ALIGNMENTS

RESULT 1

US-09-759-130B-419
; Sequence 419, Application US/09759130B
; Publication No. US2003002279A1
; GENERAL INFORMATION:
; APPLICANT: Millennium Pharmaceuticals, Inc.
; APPLICANT: McCarthy, Sean A
; APPLICANT: Fraser, Christopher C
; APPLICANT: Sharp, John D
; APPLICANT: Barnes, Thomas S
; APPLICANT: Kirt, Susan J
; APPLICANT: Mackay, Charles R
; APPLICANT: Myers, Paul S
; APPLICANT: Leiby, Kevin R
; APPLICANT: Wrighton, Nicolas
; APPLICANT: Goodearl, Andrew
; APPLICANT: Holtzman, Douglas A
; TITLE OF INVENTION: NOVEL GENES ENCODING PROTEINS HAVING PROGNOSTIC, DIAGNOSTIC, PREVENTIVE, THERAPEUTIC, AND OTHER
; TITLE OF INVENTION: USES.
; FILE REFERENCE: MPI00-5350NMIM
; CURRENT APPLICATION NUMBER: US/09/759,130B
; CURRENT FILING DATE: 2002-09-16
; PRIOR APPLICATION NUMBER: US 09/479,249
; PRIOR FILING DATE: 2000-01-07
; PRIOR APPLICATION NUMBER: US 09/559,497
; PRIOR FILING DATE: 2000-04-27
; PRIOR APPLICATION NUMBER: US 09/578,063
; PRIOR FILING DATE: 2000-05-24
; PRIOR APPLICATION NUMBER: US 09/333,159
; PRIOR FILING DATE: 1999-06-14
; PRIOR APPLICATION NUMBER: US 09/596,194
; PRIOR FILING DATE: 2000-07-14
; PRIOR APPLICATION NUMBER: US 09/342,364
; PRIOR FILING DATE: 1999-06-29
; PRIOR APPLICATION NUMBER: US 09/608,452
; PRIOR FILING DATE: 2000-06-30
; PRIOR APPLICATION NUMBER: US 09/393,996

; PRIOR FILING DATE: 1999-09-10
 ; PRIOR APPLICATION NUMBER: US 09/602,871
 ; PRIOR FILING DATE: 2000-06-23
 ; PRIOR APPLICATION NUMBER: US 09/420,707
 ; PRIOR FILING DATE: 1999-10-19
 ; NUMBER OF SEQ ID NOS: 460
 ; SOFTWARE: FastSeq for Windows Version 4.0
 ; SEQ ID NO 419
 ; LENGTH: 390
 ; TYPE: PRT
 ; ORGANISM: Homo sapiens
 ; US-09-759-130B-419

Query Match 100.0%; Score 2076; DB 11; Length 390;
 Best Local Similarity 100.0%; Pred. No. 7.3e-213;
 Matches 390; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy	1	VHMPKAVDPEAFNWI	SEIIHQGYPC	EYEVATEDGYILSVNRIPRGLVQPKTGS	RPV 60
Db	1	VHMPKAVDPEAFNWI	SEIIHQGYPC	EYEVATEDGYILSVNRIPRGLVQPKTGS	RPV 60
Qy	61	VLLQHLVGGASNI	NLNNPNSL	GFILADAGPDVWNGSRGNWSEKXKHTLSIDQ	DEFWA 120
Db	61	VLLQHLVGGASNI	NLNNPNSL	GFILADAGPDVWNGSRGNWSEKXKHTLSIDQ	DEFWA 120
Qy	121	FSYDEMARFDP	PAVINFI	LQTKQEKIYVYVGSQGTMGFIAPSTMP	ELAQIKMYFALA 180
Db	121	FSYDEMARFDP	PAVINFI	LQTKQEKIYVYVGSQGTMGFIAPSTMP	ELAQIKMYFALA 180
Qy	181	PIATVHAKSPG	KFLLPDM	MIKGLFGKKEFLYQTRFLRQLVYLCGOVIL	DDQICSNIM 240
Db	181	PIATVHAKSPG	KFLLPDM	MIKGLFGKKEFLYQTRFLRQLVYLCGOVIL	DDQICSNIM 240
Qy	241	LLGGFNTNNMNS	RSVYAAHT	LAGTSVQNIHLWSQAVNSGELRAFDWGS	ETKLEKCN 300
Db	241	LLGGFNTNNMNS	RSVYAAHT	LAGTSVQNIHLWSQAVNSGELRAFDWGS	ETKLEKCN 300
Qy	301	QPTPVRYRVD	MTVPTAM	TGGDWLSNPEDVKMLSEVTNLIYHKNIP	EWAVHVDPIWGL 360
Db	301	QPTPVRYRVD	MTVPTAM	TGGDWLSNPEDVKMLSEVTNLIYHKNIP	EWAVHVDPIWGL 360
Qy	361	DAPHRMYNEI	IHLMQQE	ETNLSQGRCEAVL 390	
Db	361	DAPHRMYNEI	IHLMQQE	ETNLSQGRCEAVL 390	

RESULT 2
 US-10-042-431-49
 ; Sequence 49, Application US/10042431
 ; Publication No. US20020182675A1
 ; GENERAL INFORMATION:
 ; APPLICANT: MCCARTHY, Sean A
 ; APPLICANT: BARNES, Thomas M
 ; APPLICANT: FRASER, Christopher C
 ; APPLICANT: SHARP, John D
 ; TITLE OF INVENTION: NOVEL GENES ENCODING PROTEINS HAVING DIAGNOSTIC,
 ; PREVENTIVE, THERAPEUTIC, AND OTHER USES
 ; FILE REFERENCE: 10147-602
 ; CURRENT APPLICATION NUMBER: US/10/042,431
 ; CURRENT FILING DATE: 2001-10-25
 ; PRIOR APPLICATION NUMBER: US 09/333,159
 ; PRIOR FILING DATE: 1999-06-14
 ; PRIOR APPLICATION NUMBER: US 09/578,063
 ; PRIOR FILING DATE: 2000-05-24
 ; NUMBER OF SEQ ID NOS: 79
 ; SOFTWARE: PatentIn Ver. 2.1
 ; SEQ ID NO 49
 ; LENGTH: 390
 ; TYPE: PRT
 ; ORGANISM: Homo sapiens
 ; US-10-042-431-49

Query Match 100.0%; Score 2076; DB 14; Length 390;

Best Local Similarity 100.0%; Pred. No. 7.3e-213;
 Matches 390; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy	1	VHMPKAVDPEAFNWI	SEIIHQGYPC	EYEVATEDGYILSVNRIPRGLVQPKTGS	RPV 60
Db	1	VHMPKAVDPEAFNWI	SEIIHQGYPC	EYEVATEDGYILSVNRIPRGLVQPKTGS	RPV 60
Qy	61	VLLQHLVGGASNI	NLNNPNSL	GFILADAGPDVWNGSRGNWSEKXKHTLSIDQ	DEFWA 120
Db	61	VLLQHLVGGASNI	NLNNPNSL	GFILADAGPDVWNGSRGNWSEKXKHTLSIDQ	DEFWA 120
Qy	121	FSYDEMARFDP	PAVINFI	LQTKQEKIYVYVGSQGTMGFIAPSTMP	ELAQIKMYFALA 180
Db	121	FSYDEMARFDP	PAVINFI	LQTKQEKIYVYVGSQGTMGFIAPSTMP	ELAQIKMYFALA 180
Qy	181	PIATVHAKSPG	KFLLPDM	MIKGLFGKKEFLYQTRFLRQLVYLCGOVIL	DDQICSNIM 240
Db	181	PIATVHAKSPG	KFLLPDM	MIKGLFGKKEFLYQTRFLRQLVYLCGOVIL	DDQICSNIM 240
Qy	241	LLGGFNTNNMNS	RSVYAAHT	LAGTSVQNIHLWSQAVNSGELRAFDWGS	ETKLEKCN 300
Db	241	LLGGFNTNNMNS	RSVYAAHT	LAGTSVQNIHLWSQAVNSGELRAFDWGS	ETKLEKCN 300
Qy	301	QPTPVRYRVD	MTVPTAM	TGGDWLSNPEDVKMLSEVTNLIYHKNIP	EWAVHVDPIWGL 360
Db	301	QPTPVRYRVD	MTVPTAM	TGGDWLSNPEDVKMLSEVTNLIYHKNIP	EWAVHVDPIWGL 360
Qy	361	DAPHRMYNEI	IHLMQQE	ETNLSQGRCEAVL 390	
Db	361	DAPHRMYNEI	IHLMQQE	ETNLSQGRCEAVL 390	

RESULT 3
 US-09-835-996A-21
 ; Sequence 21, Application US/09835996A
 ; Patent No. US20020142953A1
 ; GENERAL INFORMATION:
 ; APPLICANT: Ballinger, Dennis
 ; APPLICANT: Loeb, Debra
 ; APPLICANT: Montgomery, Julie
 ; APPLICANT: Tang, Y. Tom
 ; APPLICANT: Zhou, Ping
 ; APPLICANT: Goodrich, Ryle
 ; APPLICANT: Liu, Chenghua
 ; APPLICANT: Asundi, Vinod
 ; APPLICANT: Zhao, Qing
 ; APPLICANT: Wehrman, Tom
 ; APPLICANT: Drmanac, Radoje
 ; APPLICANT: Ren, Feiyan
 ; APPLICANT: Qian, Xiaohong
 ; APPLICANT: Wang, Dunrui
 ; TITLE OF INVENTION: MATERIALS AND METHODS RELATING TO LIPID METABOLISM
 ; FILE REFERENCE: 28110/35915A
 ; CURRENT APPLICATION NUMBER: US/09/835,996A
 ; CURRENT FILING DATE: 2001-04-16
 ; PRIOR APPLICATION NUMBER: US 60/197,137
 ; PRIOR FILING DATE: 2000-04-14
 ; PRIOR APPLICATION NUMBER: US 09/714,936
 ; PRIOR FILING DATE: 2000-11-17
 ; PRIOR APPLICATION NUMBER: US 09/667,298
 ; PRIOR FILING DATE: 2000-09-22
 ; PRIOR APPLICATION NUMBER: US 09/631,451
 ; PRIOR FILING DATE: 2000-08-03
 ; PRIOR APPLICATION NUMBER: US 09/598,042
 ; PRIOR FILING DATE: 2000-06-20
 ; NUMBER OF SEQ ID NOS: 45
 ; SOFTWARE: PatentIn version 3.0
 ; SEQ ID NO 21
 ; LENGTH: 409
 ; TYPE: PRT
 ; ORGANISM: Homo sapiens
 ; US-09-835-996A-21

Query Match 100.0%; Score 2076; DB 10; Length 409;
Best Local Similarity 100.0%; Pred. No. 7.8e-213;
Matches 390; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 VHMPTKAVDPPEAFNMISIIHQHGPCEYEYVATEDGVILSVNRIIPRGLVQPKKTGSRPV 60
DB 20 VHMPTKAVDPPEAFNMISIIHQHGPCEYEYVATEDGVILSVNRIIPRGLVQPKKTGSRPV 79
QY 61 VLLQHLVGGASNWIENLNNPNSLGFILADAGFDVVMGNSRGNMWSRKHKTLSIDQDEFWA 120
DB 80 VLLQHLVGGASNWIENLNNPNSLGFILADAGFDVVMGNSRGNMWSRKHKTLSIDQDEFWA 139
QY 121 PSYDEMARFDPVAVINFILOKTGQEKIYVYSGTTFGFIATFSTMPPELAQIKMYFALA 180
DB 140 PSYDEMARFDPVAVINFILOKTGQEKIYVYSGTTFGFIATFSTMPPELAQIKMYFALA 199
QY 181 PIATVVKHAKSPCTKFLLLPDMWIKGLFGKKEFLYQTRFLQLVYLQGVILQICSNIM 240
DB 200 PIATVVKHAKSPCTKFLLLPDMWIKGLFGKKEFLYQTRFLQLVYLQGVILQICSNIM 259
QY 241 LLLGGFNTNNMNSRASVYAAHTLAGTSVQNILHWSQAVNSGELRAFDFWGSSETKNLEKCN 300
DB 260 LLLGGFNTNNMNSRASVYAAHTLAGTSVQNILHWSQAVNSGELRAFDFWGSSETKNLEKCN 319
QY 301 QPTPVRYRVRDMTPTAMWTGGQDNLSPEDVKNLSEVTNLIYHKNIPEWAHVDFIWL 360
DB 320 QPTPVRYRVRDMTPTAMWTGGQDNLSPEDVKNLSEVTNLIYHKNIPEWAHVDFIWL 379
QY 361 DAPHRMYNEIHLMOQEBETNLSQGRCEAVL 390
DB 380 DAPHRMYNEIHLMOQEBETNLSQGRCEAVL 409

RESULT 4

US-09-759-130B-417
; Sequence 417, Application US/09759130B
; Publication No. US20030022279A1
; GENERAL INFORMATION:
; APPLICANT: Millennium Pharmaceuticals, Inc.
; APPLICANT: McCarthy, Sean A
; APPLICANT: Fraser, Christopher C
; APPLICANT: Sharp, John D
; APPLICANT: Barnes, Thomas S
; APPLICANT: KIRST, Susan J
; APPLICANT: Mackay, Charles R
; APPLICANT: Myers, Paul S
; APPLICANT: Leiby, Kevin R
; APPLICANT: Wrighton, Nicolas
; APPLICANT: Goodearl, Andrew
; APPLICANT: Holtzman, Douglas A
; TITLE OF INVENTION: PROGNOSTIC, DIAGNOSTIC, PREVENTIVE, THERAPEUTIC, AND OTHER
; FILE OF INVENTION: USES
; FILE REFERENCE: MF100-5350MIM
; CURRENT APPLICATION NUMBER: US/09759,130B
; CURRENT FILING DATE: 2002-09-16
; PRIOR APPLICATION NUMBER: US 09/479,249
; PRIOR FILING DATE: 2000-01-07
; PRIOR APPLICATION NUMBER: US 09/559,497
; PRIOR FILING DATE: 2000-04-27
; PRIOR APPLICATION NUMBER: US 09/578,063
; PRIOR FILING DATE: 2000-05-24
; PRIOR APPLICATION NUMBER: US 09/333,159
; PRIOR FILING DATE: 1999-06-14
; PRIOR APPLICATION NUMBER: US 09/596,194
; PRIOR FILING DATE: 2000-07-14
; PRIOR APPLICATION NUMBER: US 09/342,364
; PRIOR FILING DATE: 1999-06-29
; PRIOR APPLICATION NUMBER: US 09/608,452
; PRIOR FILING DATE: 2000-06-30
; PRIOR APPLICATION NUMBER: US 09/393,996
; PRIOR FILING DATE: 1999-09-10
; PRIOR APPLICATION NUMBER: US 09/602,871

; PRIOR FILING DATE: 2000-06-23
; PRIOR APPLICATION NUMBER: US 09/420,707
; PRIOR FILING DATE: 1999-10-19
; NUMBER OF SEQ ID NOS: 460
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 417
; LENGTH: 423
; TYPE: PRT
; ORGANISM: Homo sapiens
; US-09-759-130B-417

Query Match 100.0%; Score 2076; DB 11; Length 423;
Best Local Similarity 100.0%; Pred. No. 8.2e-213;
Matches 390; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 VHMPTKAVDPPEAFNMISIIHQHGPCEYEYVATEDGVILSVNRIIPRGLVQPKKTGSRPV 60
DB 34 VHMPTKAVDPPEAFNMISIIHQHGPCEYEYVATEDGVILSVNRIIPRGLVQPKKTGSRPV 93
QY 61 VLLQHLVGGASNWIENLNNPNSLGFILADAGFDVVMGNSRGNMWSRKHKTLSIDQDEFWA 120
DB 94 VLLQHLVGGASNWIENLNNPNSLGFILADAGFDVVMGNSRGNMWSRKHKTLSIDQDEFWA 153
QY 121 PSYDEMARFDPVAVINFILOKTGQEKIYVYSGTTFGFIATFSTMPPELAQIKMYFALA 180
DB 154 PSYDEMARFDPVAVINFILOKTGQEKIYVYSGTTFGFIATFSTMPPELAQIKMYFALA 213
QY 181 PIATVVKHAKSPCTKFLLLPDMWIKGLFGKKEFLYQTRFLQLVYLQGVILQICSNIM 240
DB 214 PIATVVKHAKSPCTKFLLLPDMWIKGLFGKKEFLYQTRFLQLVYLQGVILQICSNIM 273
QY 241 LLLGGFNTNNMNSRASVYAAHTLAGTSVQNILHWSQAVNSGELRAFDFWGSSETKNLEKCN 300
DB 274 LLLGGFNTNNMNSRASVYAAHTLAGTSVQNILHWSQAVNSGELRAFDFWGSSETKNLEKCN 333
QY 301 QPTPVRYRVRDMTPTAMWTGGQDNLSPEDVKNLSEVTNLIYHKNIPEWAHVDFIWL 360
DB 334 QPTPVRYRVRDMTPTAMWTGGQDNLSPEDVKNLSEVTNLIYHKNIPEWAHVDFIWL 393
QY 361 DAPHRMYNEIHLMOQEBETNLSQGRCEAVL 390
DB 394 DAPHRMYNEIHLMOQEBETNLSQGRCEAVL 423

RESULT 5

US-10-042-431-47
; Sequence 47, Application US/10042431
; Publication No. US20020182675A1
; GENERAL INFORMATION:
; APPLICANT: MCCARTHY, Sean A
; APPLICANT: BARNES, Thomas M
; APPLICANT: FRASER, Christopher C
; APPLICANT: SHARP, John D
; TITLE OF INVENTION: NOVEL GENES ENCODING PROTEINS HAVING DIAGNOSTIC,
; PREVENTIVE, THERAPEUTIC, AND OTHER USES
; FILE OF INVENTION: 10147-6U2
; CURRENT APPLICATION NUMBER: US/10/042,431
; CURRENT FILING DATE: 2001-10-25
; PRIOR APPLICATION NUMBER: US 09/333,159
; PRIOR FILING DATE: 1999-06-14
; PRIOR APPLICATION NUMBER: US 09/578,063
; PRIOR FILING DATE: 2000-05-24
; NUMBER OF SEQ ID NOS: 79
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 47
; LENGTH: 423
; TYPE: PRT
; ORGANISM: Homo sapiens
; US-10-042-431-47

Query Match 100.0%; Score 2076; DB 14; Length 423;
Best Local Similarity 100.0%; Pred. No. 8.2e-213;
Matches 390; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 VHMPTKAVDPBAPMNISEIIHQGYPCPEYEYVATEDGYILSVNRIIPRGIVQPKTGSRPV 60
 Db 34 VHMPTKAVDPBAPMNISEIIHQGYPCPEYEYVATEDGYILSVNRIIPRGIVQPKTGSRPV 93
 QY 61 VLLQGLVGGASNWSINLNNPNSLGLFADAGDVMGNSRGNAWSRKHTLSIDQDEFWA 120
 Db 94 VLLQGLVGGASNWSINLNNPNSLGLFADAGDVMGNSRGNAWSRKHTLSIDQDEFWA 153
 QY 121 FSYDENARFDLPAVINFILOKQGOEKIYVGVSGTWTGFIAPSTMPPELAQIKMYFALA 180
 Db 154 FSYDENARFDLPAVINFILOKQGOEKIYVGVSGTWTGFIAPSTMPPELAQIKMYFALA 213
 QY 181 PIATVHAKSPGKTFLLPDMNKGIFGKKEFLYQTRFLRQIVLYLCCGVILDOICSNIM 240
 Db 214 PIATVHAKSPGKTFLLPDMNKGIFGKKEFLYQTRFLRQIVLYLCCGVILDOICSNIM 273
 QY 241 LLLGGFNTNNMNSRASVVAHTLAGTSVONILHWSQAVNSGELRAFDWSETKLNKCN 300
 Db 274 LLLGGFNTNNMNSRASVVAHTLAGTSVONILHWSQAVNSGELRAFDWSETKLNKCN 333
 QY 301 QPTPVYRVRDVTPTAMWTGGQDMLSNPDEVKMLLSEVTNLIYHKNIPWAHVDPIWGL 360
 Db 334 QPTPVYRVRDVTPTAMWTGGQDMLSNPDEVKMLLSEVTNLIYHKNIPWAHVDPIWGL 393
 QY 361 DAPHRMYNEIHLMOQEEETNLSQGRCEAVL 390
 Db 394 DAPHRMYNEIHLMOQEEETNLSQGRCEAVL 423

RESULT 6
 US-09-811-825-4
 ; Sequence 4, Application US/09811825
 ; Patent No. US20020144297A1
 ; GENERAL INFORMATION:
 ; APPLICANT: YAN, Chunhua et al.
 ; TITLE OF INVENTION: ISOLATED HUMAN LIPASE PROTEINS, NUCLEIC
 ; TITLE OF INVENTION: ACID MOLECULES ENCODING HUMAN LIPASE PROTEINS, AND USES
 ; TITLE OF INVENTION: THEREOF
 ; FILE REFERENCE: CL001170
 ; CURRENT APPLICATION NUMBER: US/09/811,825
 ; CURRENT FILING DATE: 2001-03-20
 ; NUMBER OF SEQ ID NOS: 4
 ; SOFTWARE: FastSeq for Windows Version 4.0
 ; SEQ ID NO 4
 ; LENGTH: 374
 ; TYPE: PRT
 ; ORGANISM: Human
 US-09-811-825-4

Query Match 95.9%; Score 1991; DB 10; Length 374;
 Best Local Similarity 100.0%; Pred. No. 8,3e-204;
 Matches 374; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 QY 17 SEIIHQGYPCPEYEYVATEDGYILSVNRIIPRGIVQPKTGSRPVLLQHLVGGASNWS 76
 Db 1 SEIIHQGYPCPEYEYVATEDGYILSVNRIIPRGIVQPKTGSRPVLLQHLVGGASNWS 60
 QY 77 NLPNNSLGLFADAGDVMGNSRGNAWSRKHTLSIDQDEFWAFSYDEMARDLPAVIN 136
 Db 61 NLPNNSLGLFADAGDVMGNSRGNAWSRKHTLSIDQDEFWAFSYDEMARDLPAVIN 120
 QY 137 FILOKQGOEKIYVGVSGTWTGFIAPSTMPPELAQIKMYFALAPIATVYHAKSPGKFL 196
 Db 121 FILOKQGOEKIYVGVSGTWTGFIAPSTMPPELAQIKMYFALAPIATVYHAKSPGKFL 180
 QY 197 LPPDMNKGIFGKKEFLYQTRFLRQIVLYLCCGVILDOICSNIMLLGGFNTNNMNSRA 256
 Db 181 LPPDMNKGIFGKKEFLYQTRFLRQIVLYLCCGVILDOICSNIMLLGGFNTNNMNSRA 240
 QY 257 SVYAAHTLAGTSVONILHWSQAVNSGELRAFDWSETKLNKCNQPTPVYRVRDVTPT 316
 Db 241 SVYAAHTLAGTSVONILHWSQAVNSGELRAFDWSETKLNKCNQPTPVYRVRDVTPT 300

QY 317 AMWTGGQDMLSNPDEVKMLLSEVTNLIYHKNIPWAHVDPIWGLDAPHRMYNEIHLMOQ 376
 Db 301 AMWTGGQDMLSNPDEVKMLLSEVTNLIYHKNIPWAHVDPIWGLDAPHRMYNEIHLMOQ 360
 QY 377 EETNLSQGRCEAVL 390
 Db 361 EETNLSQGRCEAVL 374

RESULT 7
 US-09-811-825-2
 ; Sequence 2, Application US/09811825
 ; Patent No. US20020144297A1
 ; GENERAL INFORMATION:
 ; APPLICANT: YAN, Chunhua et al.
 ; TITLE OF INVENTION: ISOLATED HUMAN LIPASE PROTEINS, NUCLEIC
 ; TITLE OF INVENTION: ACID MOLECULES ENCODING HUMAN LIPASE PROTEINS, AND USES
 ; TITLE OF INVENTION: THEREOF
 ; FILE REFERENCE: CL001170
 ; CURRENT APPLICATION NUMBER: US/09/811,825
 ; CURRENT FILING DATE: 2001-03-20
 ; NUMBER OF SEQ ID NOS: 4
 ; SOFTWARE: FastSeq for Windows Version 4.0
 ; SEQ ID NO 2
 ; LENGTH: 398
 ; TYPE: PRT
 ; ORGANISM: Human
 US-09-811-825-2

Query Match 95.9%; Score 1991; DB 10; Length 398;
 Best Local Similarity 100.0%; Pred. No. 9,1e-204;
 Matches 374; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 QY 17 SEIIHQGYPCPEYEYVATEDGYILSVNRIIPRGIVQPKTGSRPVLLQHLVGGASNWS 76
 Db 25 SEIIHQGYPCPEYEYVATEDGYILSVNRIIPRGIVQPKTGSRPVLLQHLVGGASNWS 84
 QY 77 NLPNNSLGLFADAGDVMGNSRGNAWSRKHTLSIDQDEFWAFSYDEMARDLPAVIN 136
 Db 85 NLPNNSLGLFADAGDVMGNSRGNAWSRKHTLSIDQDEFWAFSYDEMARDLPAVIN 144
 QY 137 FILOKQGOEKIYVGVSGTWTGFIAPSTMPPELAQIKMYFALAPIATVYHAKSPGKFL 196
 Db 145 FILOKQGOEKIYVGVSGTWTGFIAPSTMPPELAQIKMYFALAPIATVYHAKSPGKFL 204
 QY 197 LPPDMNKGIFGKKEFLYQTRFLRQIVLYLCCGVILDOICSNIMLLGGFNTNNMNSRA 256
 Db 205 LPPDMNKGIFGKKEFLYQTRFLRQIVLYLCCGVILDOICSNIMLLGGFNTNNMNSRA 264
 QY 257 SVYAAHTLAGTSVONILHWSQAVNSGELRAFDWSETKLNKCNQPTPVYRVRDVTPT 316
 Db 265 SVYAAHTLAGTSVONILHWSQAVNSGELRAFDWSETKLNKCNQPTPVYRVRDVTPT 324
 QY 317 AMWTGGQDMLSNPDEVKMLLSEVTNLIYHKNIPWAHVDPIWGLDAPHRMYNEIHLMOQ 376
 Db 325 AMWTGGQDMLSNPDEVKMLLSEVTNLIYHKNIPWAHVDPIWGLDAPHRMYNEIHLMOQ 384
 QY 377 EETNLSQGRCEAVL 390
 Db 385 EETNLSQGRCEAVL 398

RESULT 8
 US-09-759-130B-411
 ; Sequence 411, Application US/09759130B
 ; Publication No. US20030022279A1
 ; GENERAL INFORMATION:
 ; APPLICANT: Millennium Pharmaceuticals, Inc.
 ; APPLICANT: McCarthy, Sean A
 ; APPLICANT: Fraser, Christopher C
 ; APPLICANT: Sharp, John D
 ; APPLICANT: Barnes, Thomas S

RESULT 11
US-10-051-874-94
; Sequence 94, Application US/10051874
; Publication No. US20040005557A1

GENERAL INFORMATION:

APPLICANT: Padigaru, Muralidhara
 APPLICANT: Alsobrook II, John P
 APPLICANT: Colman, Steven D
 APPLICANT: Spytek, Kimberly A
 APPLICANT: Boldog, Ferenc
 APPLICANT: Vernet, Corine AM
 APPLICANT: Li, Li
 APPLICANT: Shenoy, Suresh G
 APPLICANT: Casman, Stacie J
 APPLICANT: Guo, Xiaojia Sasha
 APPLICANT: Edinger, Shlomit R
 APPLICANT: MacDougall, John R
 APPLICANT: Malyankar, Uziel M
 APPLICANT: Patturajan, Meera
 APPLICANT: Shimkets, Richard A
 APPLICANT: Pena, Carol EA
 APPLICANT: Tchernev, Velizar T
 APPLICANT: Zertusen, Bryan D
 APPLICANT: Miliet, Isabelle
 APPLICANT: Miller, Charles E
 APPLICANT: Lepley, Denise M
 APPLICANT: Smithson, Glennnda
 APPLICANT: Baumgartner, Jason C
 APPLICANT: Herrman, John L
 APPLICANT: Peyman, John A
 APPLICANT: Gorman, Linda
 APPLICANT: Mezes, Peter D
 APPLICANT: Kekuda, Ramesh
 APPLICANT: Taupier Jr, Raymond J
 APPLICANT: Gerlach, Valerie
 APPLICANT: Grosse, William M
 APPLICANT: Liu, Xiaohong
 APPLICANT: Ellerman, Karen
 APPLICANT: Rothenberg, Mark
 APPLICANT: Stone, David J
 APPLICANT: Burgess, Catherine E
 TITLE OF INVENTION: PROTEINS, POLYNUCLEOTIDES ENCODING THEM AND METHODS OF
 FILE REFERENCE: 21402-245
 CURRENT APPLICATION NUMBER: US/10/051.874
 CURRENT FILING DATE: 2002-09-25
 PRIOR APPLICATION NUMBER: 60/268,595
 PRIOR FILING DATE: 2001-02-14
 PRIOR APPLICATION NUMBER: 60/325,306
 PRIOR FILING DATE: 2001-09-27
 PRIOR APPLICATION NUMBER: 60/262,587
 PRIOR FILING DATE: 2001-01-18
 PRIOR APPLICATION NUMBER: 60/272,409
 PRIOR FILING DATE: 2001-02-28
 PRIOR APPLICATION NUMBER: 60/262,454
 PRIOR FILING DATE: 2001-01-18
 PRIOR APPLICATION NUMBER: 60/276,777
 PRIOR FILING DATE: 2001-03-16
 PRIOR APPLICATION NUMBER: 60/291,672
 PRIOR FILING DATE: 2001-05-17
 PRIOR APPLICATION NUMBER: 60/330,336
 PRIOR FILING DATE: 2001-10-18
 PRIOR APPLICATION NUMBER: 60/265,530
 PRIOR FILING DATE: 2001-01-31
 PRIOR APPLICATION NUMBER: 60/261,376
 PRIOR FILING DATE: 2001-01-16
 NUMBER OF SEQ ID NOS: 269
 SOFTWARE: PatentIn Ver. 2.1
 SEQ ID NO 94
 LENGTH: 399
 TYPE: PRT
 ORGANISM: Homo sapiens
 US-10-051-874-94

Query Match 62.1%; Score 1289; DB 12; Length 399;
 Best Local Similarity 63.5%; Pred. No. 9.7e-129;
 Matches 235; Conservative 55; Mismatches 80; Indels 0; Gaps 0;

QY 7 AVDPFAFMNISEIIHQGYPCPEYEVATEDGYILSVNRIPRGLVQPKTGSPPVLLQHG 66
 DB 28 AVDPETNMNVSEIIISYWGFPSEIIVETEDGYILCLNRIPIHGRKNHSDKPKPVVFLQHG 87
 QY 67 LVGGASNIWISNLPNNISLGFILADAGFDVWNGSNGRGNWSRKHKTLSIDQDEFWAFSYDEM 126
 DB 88 LLADSSNMTNLANSLSLGFILADAGFDVWNGSNGRGNWSRKHKTLSVSDQDEFWAFSYDEM 147
 QY 127 ARFDLPVAVINFILQKTGQEKIYYVGYSGQTTMGFTAFSTMPDLAQIKMYFALAPIATVK 186
 DB 148 AKYDLPASINFILNKTGQEQVYVGHSGQTTGFTAFSQIPELAKRIKNFFALGFVASVA 207
 QY 187 HAKSPGTGTELLPDMMIKGLFGKKEFLYQTRFLRQLVYLCOVILDOICSNIMLLGGF 246
 DB 208 FCTSPMAKLGRLPDHLIKDLFGDKFLQSAFLKWLGVTHVCHVILKELCGNLCFLLCGF 267
 QY 247 NTNNMMSRASYAAHTLAGTSVQNIHWSQAVNGSELRAFDWGETKNLEKCNQPTPVR 306
 DB 268 NERNLMSRVDVYTTSTPAGTSVQNMLEWSQAVKPKQAFDQWGSASAKNYPHYNSYPT 327
 QY 307 YRVDMTVPATAMTGGQDWLSNPFEDVKMLLSVTNLIYHKNIPFWAHVDFTWGLDAPHRM 366
 DB 328 YNVKDMLVPTAVWSGHDMLADVVDVNIILLTQITNLVFHESIPWEHLDFTINGLDAPWRL 387
 QY 367 YNEIHLMOQ 376
 DB 388 YNKIINLMEK 397

RESULT 12
 US-10-042-431-41
 ; Sequence 41, Application US/10042431
 ; Publication NO. US20020182675A1
 ; GENERAL INFORMATION:
 ; APPLICANT: MCCARTHY, Sean A
 ; APPLICANT: BARNES, Thomas M
 ; APPLICANT: FRASER, Christopher C
 ; APPLICANT: SHARP, John D
 ; TITLE OF INVENTION: NOVEL GENES ENCODING PROTEINS HAVING DIAGNOSTIC,
 ; FILE REFERENCE: 10147-6U2
 ; CURRENT APPLICATION NUMBER: US/10/042,431
 ; CURRENT FILING DATE: 2001-10-25
 ; PRIOR APPLICATION NUMBER: US 09/333,159
 ; PRIOR FILING DATE: 1999-06-14
 ; PRIOR APPLICATION NUMBER: US 09/578,063
 ; PRIOR FILING DATE: 2000-05-24
 ; NUMBER OF SEQ ID NOS: 79
 ; SOFTWARE: PatentIn Ver. 2.1
 ; SEQ ID NO 41
 ; LENGTH: 399
 ; TYPE: PRT
 ; ORGANISM: Homo sapiens
 US-10-042-431-41

Query Match 62.1%; Score 1289; DB 14; Length 399;
 Best Local Similarity 63.5%; Pred. No. 9.7e-129;
 Matches 235; Conservative 55; Mismatches 80; Indels 0; Gaps 0;

QY 7 AVDPFAFMNISEIIHQGYPCPEYEVATEDGYILSVNRIPRGLVQPKTGSPPVLLQHG 66
 DB 28 AVDPETNMNVSEIIISYWGFPSEIIVETEDGYILCLNRIPIHGRKNHSDKPKPVVFLQHG 87
 QY 67 LVGGASNIWISNLPNNISLGFILADAGFDVWNGSNGRGNWSRKHKTLSIDQDEFWAFSYDEM 126
 DB 88 LLADSSNMTNLANSLSLGFILADAGFDVWNGSNGRGNWSRKHKTLSVSDQDEFWAFSYDEM 147
 QY 127 ARFDLPVAVINFILQKTGQEKIYYVGYSGQTTMGFTAFSTMPDLAQIKMYFALAPIATVK 186
 DB 148 AKYDLPASINFILNKTGQEQVYVGHSGQTTGFTAFSQIPELAKRIKNFFALGFVASVA 207
 QY 187 HAKSPGTGTELLPDMMIKGLFGKKEFLYQTRFLRQLVYLCOVILDOICSNIMLLGGF 246

```

Db      208  FCTSPMAKLGRLPDHLIKDLFGDKFPLQSAFLKMLGTHVTHVILKELCGNLCFLLCGF 267
Qy      247  NTNNMMSRASVYAAHTLACTSVONILHWSQAVNSGELAPDWSGSETKMLEKCNQPTVR 306
Db      268  NERNLNMSRDVVTTHSPAGTSVQNNMLHWSQAVKQFQAFDWSGSAKXNYFHYNSYPT 327
Qy      307  YRVDMTVPMTAMTGGODMLSNPDYKMLISEVTNLIYHKNIPWAHVDPIFGLDAPHRM 366
Db      328  YNVKMDLVPVAVSGGHDWLDVVDVNNILITQITNLVFHESIPWEHDLFIINGLDAPWRL 387
Qy      367  YNEIHLMOQ 376
Db      388  YNKIINLMRK 397

```

```

RESULT 13
US-10-003-302-4
; Sequence 4, Application US/10003302
; Publication No. US20020142435A1
; GENERAL INFORMATION:
; APPLICANT: MERKULOV, Gennady et al.
; TITLE OF INVENTION: ISOLATED HUMAN LIPASE PROTEINS, NUCLEIC
; TITLE OF INVENTION: ACID MOLECULES ENCODING HUMAN LIPASE PROTEINS, AND USES
; TITLE OF INVENTION: THEREOF
; FILE REFERENCE: CL001186DIV
; CURRENT APPLICATION NUMBER: US/10/003,302
; CURRENT FILING DATE: 2001-12-06
; NUMBER OF SEQ ID NOS: 4
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 4
; LENGTH: 392
; TYPE: PRT
; ORGANISM: Homo sapiens
US-10-003-302-4

```

```

Query Match      61.8%; Score 1282; DB 14; Length 392;
Best Local Similarity 63.8%; Pred. No. 5.3e-128;
Matches 234; Conservative 53; Mismatches 80; Indels 0; Gaps 0;

Qy      7  AVDPAPFNISEIIHQGYPCPEYEVEDGYILSVNRIPRGLVQPKTGSRRVLLQHG 66
Db      26  AVDPETNMYSEIISYWGFPSEYLVETEDGYILCLNRIPIHGRKNHSDGKPEWVFLQHG 85
Qy      67  LVGASNWSNLNPNLSGLFADAGFDVWNGNSRGNWSRKHKTLSIDQDEFWAFSYDEM 126
Db      86  LLADSSNWTNLANSLSGLFADAGFDVWNGNSRGNTWSEKHTLSVSDQDEFWAFSYDEM 145
Qy      127  ARFDLPVAVNFILOKTQGEKIYVYVGSQGTMTGFIAPSTWPELAQKIMYFALAPIATVK 186
Db      146  AKYDLPASINFILNKTQGEQVYVGHSGQTTIGFIAPSOIPELAKIKRFFALGPVASYA 205
Qy      187  HAKSPGPKFLLPDMMIKGLFGKKEFLYQTRFLRQLVYLCOGVILDOICSNIMLLGFG 246
Db      206  FCTSPMAKLGRLPDHLIKDLFGDKFPLQSAFLKMLGTHVTHVILKELCGNLCFLLCGF 265
Qy      247  NTNNMMSRASVYAAHTLACTSVONILHWSQAVNSGELAPDWSGSETKMLEKCNQPTVR 306
Db      266  NERNLNMSRDVVTTHSPAGTSVQNNMLHWSQAVKQFQAFDWSGSAKXNYFHYNSYPT 325
Qy      307  YRVDMTVPMTAMTGGODMLSNPDYKMLISEVTNLIYHKNIPWAHVDPIFGLDAPHRM 366
Db      326  YNVKMDLVPVAVSGGHDWLDVVDVNNILITQITNLVFHESIPWEHDLFIINGLDAPWRL 385
Qy      367  YNEIHL 373
Db      386  YNKIINL 392

```

```

RESULT 14
US-09-759-130B-420
; Sequence 420, Application US/09759130B
; Publication No. US2003002279A1

```

```

; GENERAL INFORMATION:
; APPLICANT: Millennium Pharmaceuticals, Inc.
; APPLICANT: McCarthy, Sean A
; APPLICANT: Fraser, Christopher C
; APPLICANT: Sharp, John D
; APPLICANT: Baries, Thomas S
; APPLICANT: Kirt, Susan J
; APPLICANT: Mackay, Charles R
; APPLICANT: Myers, Paul S
; APPLICANT: Leiby, Kevin R
; APPLICANT: Wrighton, Nicolas
; APPLICANT: Goodearl, Andrew
; APPLICANT: Holtzman, Douglas A
; TITLE OF INVENTION: NOVEL GENES ENCODING PROTEINS HAVING
; TITLE OF INVENTION: PROGNOSTIC, DIAGNOSTIC, PREVENTIVE, THERAPEUTIC, AND OTHER
; TITLE OF INVENTION: USES
; FILE REFERENCE: MP100-535OMNIM
; CURRENT APPLICATION NUMBER: US/09/759,130B
; CURRENT FILING DATE: 2002-09-16
; PRIOR APPLICATION NUMBER: US 09/479,249
; PRIOR FILING DATE: 2000-01-07
; PRIOR APPLICATION NUMBER: US 09/559,497
; PRIOR FILING DATE: 2000-04-27
; PRIOR APPLICATION NUMBER: US 09/578,063
; PRIOR FILING DATE: 2000-05-24
; PRIOR APPLICATION NUMBER: US 09/333,159
; PRIOR FILING DATE: 1999-06-14
; PRIOR APPLICATION NUMBER: US 09/596,194
; PRIOR FILING DATE: 2000-07-14
; PRIOR APPLICATION NUMBER: US 09/342,364
; PRIOR FILING DATE: 1999-06-29
; PRIOR APPLICATION NUMBER: US 09/608,452
; PRIOR FILING DATE: 2000-06-30
; PRIOR APPLICATION NUMBER: US 09/393,996
; PRIOR FILING DATE: 1999-09-10
; PRIOR APPLICATION NUMBER: US 09/602,871
; PRIOR FILING DATE: 2000-06-23
; PRIOR APPLICATION NUMBER: US 09/420,707
; PRIOR FILING DATE: 1999-10-19
; NUMBER OF SEQ ID NOS: 460
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 420
; LENGTH: 221
; TYPE: PRT
; ORGANISM: Homo sapiens
US-09-759-130B-420

```

```

Query Match      56.1%; Score 1165; DB 11; Length 221;
Best Local Similarity 100.0%; Pred. No. 7.1e-116;
Matches 221; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy      1  VHMPTKAVDPEAFNISEIIHQGYPCPEYEVEDGYILSVNRIPRGLVQPKTGSRRV 60
Db      1  VHMPTKAVDPEAFNISEIIHQGYPCPEYEVEDGYILSVNRIPRGLVQPKTGSRRV 60
Qy      61  VLLQHLVGGASNWSNLNPNLSGLFADAGFDVWNGNSRGNWSRKHKTLSIDQDEFWA 120
Db      61  VLLQHLVGGASNWSNLNPNLSGLFADAGFDVWNGNSRGNWSRKHKTLSIDQDEFWA 120
Qy      121  FSYDEMARFDLPVAVNFILOKTQGEKIYVYVGSQGTMTGFIAPSTWPELAQKIMYFALA 180
Db      121  FSYDEMARFDLPVAVNFILOKTQGEKIYVYVGSQGTMTGFIAPSTWPELAQKIMYFALA 180
Qy      181  PIATVXKASPGTKFLLPDMMIKGLFGKKEFLYQTRFLRQ 221
Db      181  PIATVXKASPGTKFLLPDMMIKGLFGKKEFLYQTRFLRQ 221

```

```

RESULT 15
US-10-042-431-50
; Sequence 50, Application US/10042431
; Publication No. US20020182675A1
; GENERAL INFORMATION:

```

; APPLICANT: MCCARTHY, Sean A
 ; APPLICANT: BARNES, Thomas M
 ; APPLICANT: FRASER, Christopher C
 ; APPLICANT: SHARP, John D
 ; TITLE OF INVENTION: NOVEL GENES ENCODING PROTEINS HAVING DIAGNOSTIC,
 ; TITLE OF INVENTION: PREVENTIVE, THERAPEUTIC, AND OTHER USES
 ; FILE REFERENCE: 10147-6U2
 ; CURRENT APPLICATION NUMBER: US/10/042.431
 ; PRIOR FILING DATE: 2001-10-25
 ; PRIOR FILING DATE: 1999-06-14
 ; PRIOR FILING DATE: 2000-05-24
 ; NUMBER OF SEQ ID NOS: 79
 ; SOFTWARE: Patentin Ver. 2.1
 ; SEQ ID NO 50
 ; LENGTH: 221
 ; TYPE: PRT
 ; ORGANISM: Homo sapiens
 US-10-042-431-50

Query Match 56.1%; Score 1165; DB 14; Length 221;
 Best Local Similarity 100.0%; Pred. No. 7.1e-116;
 Matches 221; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 Qy 1 VHMPTKAVDPEAFNISEIIHQGYPCPEYEYEVATEDGYILSVNRI PRGLVQPKKTSRPPV 60
 Db 1 VHMPTKAVDPEAFNISEIIHQGYPCPEYEYEVATEDGYILSVNRI PRGLVQPKKTSRPPV 60
 Qy 61 VLLQHLVGGASNNWISNLPNNLSGLFILADAGFDVWMCNSRGNAWSRKHKHTLSIDQDEFWA 120
 Db 61 VLLQHLVGGASNNWISNLPNNLSGLFILADAGFDVWMCNSRGNAWSRKHKHTLSIDQDEFWA 120
 Qy 121 PSYDEMARFDLPVAVINFILOKTGOEKIYYGYSGQITMGFIAPSTWPELAQKIKMYFALA 180
 Db 121 PSYDEMARFDLPVAVINFILOKTGOEKIYYGYSGQITMGFIAPSTWPELAQKIKMYFALA 180
 Qy 181 PIATVGHAKSPGTFKLLLPDMWIKGLFGKKEFLYQTRFLRQ 221
 Db 181 PIATVGHAKSPGTFKLLLPDMWIKGLFGKKEFLYQTRFLRQ 221

Search completed: February 19, 2004, 14:34:26
 Job time : 28.3838 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2004 Compugen Ltd.

OM protein - protein search, using sw model

Run on: February 19, 2004, 14:17:49 ; Search time 13.4317 Seconds
(without alignments)
2792.326 Million cell updates/sec

Title: US-10-042-431-49

Perfect score: 2076
Sequence: 1 VHMPTKAVDPEAFNMISEII.....IHLMOQBETNLSQRCRAVL 390

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 283308 seqs, 96168682 residues

Total number of hits satisfying chosen parameters: 283308

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database :

PIR 76:*
1: PIR1.*
2: PIR2.*
3: PIR3.*
4: PIR4.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Match	Length	DB ID	Description
1	1289	62.1	399	2 S41408	lysosomal acid lip
2	1289	62.1	399	2 G01416	lysosomal acid lip
3	1161	55.9	395	1 L1RTT	triacylglycerol li
4	1125	54.2	397	1 J4017	triacylglycerol li
5	1109	53.4	398	2 S07145	triacylglycerol li
6	747	36.0	411	2 T22290	hypothetical prote
7	744.5	35.9	403	2 T33198	hypothetical prote
8	727	35.0	405	2 T22675	hypothetical prote
9	699	33.7	405	2 H88930	protein R11G11.14
10	691.5	33.3	411	2 G89074	protein K04A8.5 [i
11	651.5	31.4	426	2 T30480	hypothetical prote
12	574	27.6	559	2 J70949	egg-specific prote
13	491	23.7	443	2 T39540	triglyceride lipas
14	478.5	23.0	1585	2 T31611	hypothetical prote
15	475	22.9	445	2 T41053	triglyceride lipas
16	445	21.4	344	2 E84526	probable lysosomal
17	422.5	20.4	548	2 S37969	probable triacylg
18	403	19.4	460	2 T39443	probable triacylg
19	397.5	19.1	413	2 T33170	probable triacylg
20	294	14.2	431	2 D86318	protein F5H18.6 [i
21	282	13.6	538	2 S64842	probable membrane
22	271.5	13.1	573	2 S64754	protein lipase F2p
23	232.5	11.2	509	2 G96766	lipase, pectinase
24	149	7.2	62	2 S59904	probable esterase/
25	143	6.9	336	2 P83425	conserved hypothet
26	113	5.4	412	2 B72391	triacylglycerol li
27	112.5	5.4	40	2 S29539	hypothetical prote
28	111.5	5.4	987	2 A64474	lysophospholipase
29	107	5.2	281	2 A11912	

ALIGNMENTS

RESULT 1

S41408
lysosomal acid lipase (EC 3.1.1.1.-) / sterol esterase (EC 3.1.1.13) precursor - human
C;Species: Homo sapiens (man)
C;Date: 13-Jan-1995 #sequence revision 13-Jan-1995 #text_change 18-Jun-1999
C;Accession: S41408; A39315; S47187
R;Aneis, D.; Merkel, M.; Eckerskorn, C.; Greten, H.
Eur. J. Biochem. 219, 905-914, 1994
A;Title: Purification, characterization and molecular cloning of human hepatic lysosomal
A;Reference number: S41408; MUID:94155897; PMID:8112342
A;Status: preliminary
A;Molecule type: mRNA
A;Residues: 1-399 <AME>
A;Cross-references: EMBL:X76488; NID:G4344305; PIDN:CAA54026.1; PID:G4344306
R;Anderson, R.A.; Sando, G.N.
J. Biol. Chem. 266, 22479-22484, 1991
A;Title: Cloning and expression of cDNA encoding human lysosomal acid lipase/cholesteryl
A;Reference number: A39315; MUID:92042192; PMID:1718995
A;Accession: A39315
A;Status: preliminary
A;Molecule type: mRNA
A;Residues: 1-15, 'P', 17-399 <AND>
A;Cross-references: GB:M74775; NID:G187151; PIDN:AAA59519.1; PID:G187152
R;Du, H.; Gregory, G.A.
submitted to the EMBL Data Library, April 1994
A;Description: Structural conservation of putative functional motifs between mouse and h
A;Reference number: S47187
A;Accession: S47187
A;Status: preliminary
A;Molecule type: mRNA
A;Residues: 1-22, 'R', 24-399 <DUH>
A;Cross-references: EMBL:231690; NID:G506430; PIDN:CAA83495.1; PID:G506431
C;Genetics:
A;Gene: GDB:LIPA
A;Cross-references: GDB:120153; OMIM:278000
A;Map position: 10q24-10q25
C;Superfamily: triacylglycerol lipase, lingual
C;Keywords: carboxylic ester hydrolase; glycoprotein

Query Match 62.1%; Score 1289; DB 2; Length 399;

Best Local Similarity 63.5%; Pred. No. 8.9e-107;

Matches 235; Conservative 55; Mismatches 80; Indels 0; Gaps 0;

Qy 7 AVDPEAFNMISEIIHQGYCFEEYEVATEDGYILSVNRIPIRGVLQPKTKGSPVYLLQHG 66

Db 28 AVDPEAFNMISEIIHQGYCFEEYEVATEDGYILSVNRIPIRGVLQPKTKGSPVYLLQHG 87

Qy 67 LVGGASNIWSNLPNNSLGLFILADAGFDVWNGSRGNWSRKHTLSIQDDEFWAFSYDEM 126

Db 88 LLADSSNWTNLSGLFILADAGFDVWNGSRGNWSRKHTLSIQDDEFWAFSYDEM 147

QY 127 ARFDLPAVINFILOKQGEKIYVYVGSQGTMTGFIAPFTWPELAOKIKMYFALAPATYK 186
 Db 148 AKYDLPASINFILOKQGEKIYVYVGSQGTMTGFIAPFTWPELAOKIKMYFALGPVAVSA 207
 QY 187 HAKSPGKFKLLPDMNKGKFGKKEFLYQTRFRLQVILVCGVILDOICSNMILLGQF 246
 Db 208 FCTSPMAKGLPDLHLKDLFGDKFLPQSAFLKWLGVTHVCTVILKELCGNLCFLGQF 267
 QY 247 NTNNMNSRASVYAAHTLAGTSVQNLHWSQAVNSGELRAFDWGSSTKLEKCNQPTPYR 306
 Db 268 NERNLNMSRDVVTTHSPAGTSVQNLHWSQAVNSGELRAFDWGSSTKLEKCNQPTPYR 327
 QY 307 YRVDMVTPTAMTGGQDWLNSPDEVKMLSEVTLNLYHKNIPPEWAVDFPGLDAPHRM 366
 Db 328 YNVKMDLVPVAVWSGGHDWLDVYDVNILLTQITNLVPHESIPWEHDLFPGLDAPWRL 387
 QY 367 YNEIHLMOQ 376
 Db 388 YNKIINLMRK 397

RESULT 2

G01416
 lysosomal acid lipase - human
 C:Species: Homo sapiens (man)
 C:Date: 21-Dec-1996 #sequence_revision 06-Jun-1997 #text_change 18-Jun-1999
 C:Accession: G01416
 R:Du, H.

submitted to the EMBL Data Library, April 1994

A:Reference number: G06919
 A:Accession: G01416
 A:Status: preliminary; translated from GB/EMBL/DBJ
 A:Molecule type: mRNA
 A:Residues: 1-399 <DUX>
 A:Cross-references: EMBL:U08464; NID:G505052; PIDN:AA60328.1; PID:G505053
 C:Superfamily: triacylglycerol lipase, lingual
 C:Keywords: Glycoprotein

Query Match 62.1%; Score 1289; DB 2; Length 399;
 Best Local Similarity 63.5%; Pred. No. 8.9e-107;
 Matches 235; Conservative 55; Mismatches 80; Indels 0; Gaps 0;

QY 7 AVDEAFMNISETIIHQGYPCREYEVATEDGYILSVNRIPRGIVQPKTGRSPVLLQHG 56
 Db 28 AVDEPMNVSEIIISYGFPSSEYLVETEDGYILCLNRIPHGRKNHSDGPKPVEVLQHG 87
 QY 67 LVGASNWIENLPNNSLGFILADAGFDVWNGSRGNAMSRKHKTLSIDODEFWAFSYDEM 126
 Db 88 LLADSSNWVNLANSLSLGFILADAGFDVWNGSRGNAMSRKHKTLSVSDQEFWAFSYDEM 147
 QY 127 ARFDLPAVINFILOKQGEKIYVYVGSQGTMTGFIAPFTWPELAOKIKMYFALAPATYK 186
 Db 148 AKYDLPASINFILOKQGEKIYVYVGSQGTMTGFIAPFTWPELAOKIKMYFALGPVAVSA 207
 QY 187 HAKSPGKFKLLPDMNKGKFGKKEFLYQTRFRLQVILVCGVILDOICSNMILLGQF 246
 Db 208 FCTSPMAKGLPDLHLKDLFGDKFLPQSAFLKWLGVTHVCTVILKELCGNLCFLGQF 267
 QY 247 NTNNMNSRASVYAAHTLAGTSVQNLHWSQAVNSGELRAFDWGSSTKLEKCNQPTPYR 306
 Db 268 NERNLNMSRDVVTTHSPAGTSVQNLHWSQAVNSGELRAFDWGSSTKLEKCNQPTPYR 327
 QY 307 YRVDMVTPTAMTGGQDWLNSPDEVKMLSEVTLNLYHKNIPPEWAVDFPGLDAPHRM 366
 Db 328 YNVKMDLVPVAVWSGGHDWLDVYDVNILLTQITNLVPHESIPWEHDLFPGLDAPWRL 387
 QY 367 YNEIHLMOQ 376
 Db 388 YNKIINLMRK 397

RESULT 3

LIRIT

triacylglycerol lipase (EC 3.1.1.3) precursor, lingual - rat
 C:Species: Rattus norvegicus (Norway rat)
 C:Date: 30-Sep-1987 #sequence_revision 30-Sep-1987 #text_change 11-Jun-1999
 C:Accession: A23045
 R:Doeherty, A.J.P.; Bodmer, M.W.; Angal, S.; Verger, R.; Riviere, C.; Lowe, P.A.; Lyons, Nucleic Acids Res. 13, 1891-1903, 1985
 A:Title: Molecular cloning and nucleotide sequence of rat lingual lipase cDNA.
 A:Reference number: A23045; MUID:85215587; PMID:3839077
 A:Accession: A23045
 A:Molecule type: mRNA
 A:Residues: 1-395 <DOC>
 A:Cross-references: GB:X02309; NID:G56595; PIDN:CAA26179.1; PID:G56596
 A:Experimental source: seran Sprague-Dawley
 A:Note: The partial sequence of the mature protein from a different, unspecified strain
 C:Comment: This acid-stable lipase is secreted by the serous (von Ebner's) glands at the
 C:Superfamily: triacylglycerol lipase, lingual
 C:Keywords: carboxylic ester hydrolase; glycoprotein; lipid digestion; saliva; serous gl
 F:1-18/Domain: signal sequence #status predicted <SIG>
 F:19-395/Product: triacylglycerol lipase, lingual #status predicted <MPT>
 F:33,68,98,184,270/Binding site: carbohydrate (Asn) (covalent) #status predicted

Query Match 55.9%; Score 1161; DB 1; Length 395;
 Best Local Similarity 57.7%; Pred. No. 2.2e-95;
 Matches 213; Conservative 63; Mismatches 93; Indels 0; Gaps 0;

QY 9 DPEAFMNISETIIHQGYPCREYEVATEDGYILSVNRIPRGIVQPKTGRSPVLLQHG 68
 Db 27 NPEANMISQMTIYWGYPCEYEVATEDGYILGYRIPHGKNSENIGKRPVYLQGLI 86
 QY 69 CGASNWIENLPNNSLGFILADAGFDVWNGSRGNAMSRKHKTLSIDODEFWAFSYDEM 128
 Db 87 ASATNWIENLPNNSLGFILADAGFDVWNGSRGNAMSRKHKTLSIDODEFWAFSYDEM 146
 QY 129 FDLPAVINFILOKQGEKIYVYVGSQGTMTGFIAPFTWPELAOKIKMYFALAPATYK 188
 Db 147 YDLPATNFILOKQGEKIYVYVGSQGTMTGFIAPFTWPELAOKIKMYFALAPATYK 206
 QY 189 KSPGKFKLLPDMNKGKFGKKEFLYQTRFRLQVILVCGVILDOICSNMILLGQFNT 248
 Db 207 QSLPKKISFIPTFLFKLFGKKEFLYQTRFRLQVILVCGVILDOICSNMILLGQFNT 266
 QY 249 NNMNMSRASVYAAHTLAGTSVQNLHWSQAVNSGELRAFDWGSSTKLEKCNQPTPYR 308
 Db 267 KNLNMSRDVVTTHSPAGTSVQNLHWSQAVNSGELRAFDWGSSTKLEKCNQPTPYR 326
 QY 309 VRDMVTPTAMTGGQDWLNSPDEVKMLSEVTLNLYHKNIPPEWAVDFPGLDAPHRM 368
 Db 327 VSAMTVPVAVWSGGHDWLDVYDVNILLTQITNLVPHESIPWEHDLFPGLDAPHRM 386
 QY 369 EIIHLMQOE 377
 Db 387 EIMSMMAED 395

RESULT 4

JC4017
 triacylglycerol lipase (EC 3.1.1.3) PGF precursor - bovine
 N:Alternate names: pregastric esterase
 C:Species: Bos primigenius taurus (cattle)
 C:Date: 10-Sep-1999 #sequence_revision 10-Sep-1999 #text_change 10-Sep-1999
 C:Accession: JC4017; S64678
 R:Timmermans, M.Y.J.; Teuchy, H.; Kupers, L.P.M.
 Gene 147, 259-262, 1994
 A:Title: The cDNA sequence encoding bovine pregastric esterase.
 A:Reference number: JC4017; MUID:95011625; PMID:7926811
 A:Accession: JC4017
 A:Molecule type: mRNA
 A:Residues: 1-397 <TIM>
 A:Cross-references: GB:L26319; NID:G600756; PIDN:AAA57037.1; PID:G600757
 A:Experimental source: tongue
 R:Timmermans, M.Y.J.; Reekmans, G.; Teuchy, H.J.H.; Kupers, L.P.M.
 Biochem. J. 314, 931-936, 1996

A:Title: Inhibition studies on calf pregastric esterase: the enzyme has no functional ti

A;Reference number: S64678; MUID:96177869; PMID:8615791

A;Accession: S64678

A;Molecule type: protein

A;Residues: 24-407248-253 <N>

C;Comment: Pregastric esterase is a major fat-digesting enzyme.

C;Genetics:

A;Gene: pge

C;Superfamily: triacylglycerol lipase, lingual

C;Keywords: blocked amino end; carboxylic ester hydrolase; glycoprotein

F;1-19/Domain: signal sequence #status predicted <SIG>

F;20-397/Product: pregastric esterase #status predicted <NAT>

F;33-270/Binding site: carbohydrate (Asn) (covalent) #status predicted

F;171/Active site: Ser #status predicted

Query Match 54.2%; Score 1125; DB 1; Length 397;
Best Local Similarity 55.8%; Pred. No. 3.5e-92;
Matches 208; Conservative 64; Mismatches 99; Indels 2; Gaps 2;

QY 7 AVDPAPNISEIIOHQYPCBEYEVATEDGYILSVNRIIPGLVOPKKTGSRPVVLLQHG 66

DB 25 AKNPEASNVVSQMSIYWGYPSEMHKVIADGYILQVYRIPHGKNNANHLGQRPVVFLQHG 84

QY 67 LVGGASNIISLNPNSLGFILADAGFDVWNGSRGNASRKHKTLISIDQDEFWAFSDYDEM 126

DB 85 LLGSATNLSLNPNSLGFILADAGFDVWNGSRGNWQAHELYSPDSPFWAFSDYDEM 144

QY 127 ARFLPAPNIFLOKTGQEKIYYGYSGQTMTGFIATFMTPELAQIKMYFALAPIATYK 186

DB 145 AYEYDLPTIDFLRRTGQKGLHYVGHSGQTMTGFIATFMTPELAQIKMYFALAPIATYK 204

QY 187 HAKSPGTFLPLPDMWIKGLGKKEFLYQTRFLQ-1VYILCGQVILDOICSNIMLLGG 245

DB 205 YKSLFNKALIPHELFKIIIGDKMF-YPHFLQGLGVEMCSRETLVLCKNALFAITG 263

QY 246 FNTNMNMSRASYAAHTLAGTSVQNIILHWSQAVNSGELRAFDWGSSEKLEKCNQTPV 305

DB 264 VDNKNFNSRLDYIAHNPAGTSVQNTLHWRQVSKGKQFADWGAQVQNLHMYHQTPP 323

QY 306 RYRVDMTPVPMWVGQDWSNPEDVKMLSEVTLNLYHKNIPEWAVHVDINGLDAPHR 365

DB 324 IYNTAMNVPYAVWSADNDLADPDVDFLLSKSLNLYHKNIPEWAVHVDINGLDAPHR 383

QY 366 MYNEIHLMOQEE 378

DB 384 VYNEIVSLMAEDK 396

RESULT 5

S07145

triacylglycerol lipase (EC 3.1.1.3) precursor, gastric - human

C;Species: Homo sapiens (man)

C;Date: 29-Jan-1993 #sequence revision 29-Jan-1993 #text_change 18-Jun-1999

C;Accession: S07145; S27102; S04942

R;Bodmer, M.W.; Angal, S.; Yarranton, G.T.; Harris, T.J.R.; Lyons, A.; King, D.J.; Piero

Biochem. Biophys. Acta 909, 237-244, 1987

A;Title: Molecular cloning of a human gastric lipase and expression of the enzyme in yeast

A;Reference number: S07145; MUID:87299724; PMID:3304425

A;Accession: S07145

A;Molecule type: mRNA

A;Residues: 1-398 <B01>

A;Cross-references: EMBL:X05997; NID:g31771; PIDN:CAA29413.1; PID:g758063

A;Accession: S27102

A;Molecule type: protein

A;Residues: 20-24,'X',26,'X',28-33,'X',35-42,'X',44,'X',46-48,'X',50,'X',52,'X',54-56 <E

A;Note: it is uncertain whether Met-1 or Met-7 is the initiator

R;Bernbaeck, S.; Blaackberg, L.

Eur. J. Biochem. 182, 495-499, 1989

A;Title: Human gastric lipase. The N-terminal tetrapeptide is essential for lipid binding

A;Reference number: S04942; MUID:89325292; PMID:2753032

A;Accession: S04942

A;Molecule type: protein

A;Residues: 20-45 <BER>

C;Superfamily: triacylglycerol lipase, lingual

C;Keywords: carboxylic ester hydrolase; extracellular protein; glycoprotein; lipid digesti
F;1-19/Domain: signal sequence #status predicted <SIG>
F;20-398/Product: triacylglycerol lipase, gastric #status experimental <NAT>
F;34,99,185,271/Binding site: carbohydrate (Asn) (covalent) #status predicted

Query Match 53.4%; Score 1109; DB 2; Length 398;

Best Local Similarity 54.7%; Pred. No. 9.5e-91;

Matches 202; Conservative 69; Mismatches 98; Indels 0; Gaps 0;

QY 10 PEAPNISEIIOHQYPCBEYEVATEDGYILSVNRIIPGLVOPKKTGSRPVVLLQHLVG 69

DB 29 PEVTNNISQMITWGYNPEEYEVATEDGYILEVNRIPYKKNNGTGORPVVFLQHLGA 88

QY 70 GASNIISLNPNSLGFILADAGFDVWNGSRGNASRKHKTLISIDQDEFWAFSDYEMARF 129

DB 89 SATNLSLNPNSLGFILADAGFDVWNGSRGNWARRNLYSPDSVEFWAFSDYEMAKY 148

QY 130 DLPAPNIFLOKTGQEKIYYGYSGQTMTGFIATFMTPELAQIKMYFALAPIATYKAK 189

DB 149 DLPATIDFIVKTKGQKGLHYVGHSGQTMTGFIATFMTPELAQIKMYFALAPIATYKTK 208

QY 190 SPOTKELLPLPDMWIKGLGKKEFLYQTRFLQ-1VYILCGQVILDOICSNIMLLGFTN 249

DB 209 SLINKLRFVPSLFXEFGDKLIFYPHNPFDFQLATEVCSREMLNLLCSNALFIICGFSK 268

QY 250 NPMMSRASYAAHTLAGTSVQNIILHWSQAVNSGELRAFDWGSSEKLEKCNQTPVRYV 309

DB 269 NFNTSLDLYLHNPAGTSVQNNFHWQVSKGKQAYDWGSPVQNRHMYDQSPPYNV 328

QY 310 RDMVTPTAMTGGQDWSNPEDVKMLSEVTLNLYHKNIPEWAVHVDINGLDAPHRMYNE 369

DB 329 TAMNVPYAVWSADNDLADPDVDFLLSKSLNLYHKNIPEWAVHVDINGLDAPHRMYNE 388

QY 370 IHLMOQEE 378

DB 389 IVSMISEDK 397

RESULT 6

T22290

hypothetical protein F46B6.8 - Caenorhabditis elegans

C;Species: Caenorhabditis elegans

C;Date: 15-Oct-1999 #sequence_revision 15-Oct-1999 #text_change 21-Jan-2000

C;Accession: T22290

R;McMurray, A.

submitted to the EMBL Data Library, April 1996

A;Reference number: Z19542

A;Accession: T22290

A;Status: preliminary; translated from GB/EMBL/DBJ

A;Molecule type: DNA

A;Residues: 1-411 <WIL>

A;Cross-references: EMBL:270780; PIDN:CAA94824.1; GSPDB:GN000023; CESP:F46B6.8

A;Experimental source: clone F46B6

C;Genetics:

A;Gene: CESP:F46B6.8

A;Map position: 5

A;Introns: 35/3; 69/1; 104/1; 226/3

C;Superfamily: triacylglycerol lipase, lingual

Query Match

Best Local Similarity 36.0%; Score 747; DB 2; Length 411;

Matches 159; Conservative 66; Mismatches 130; Indels 26; Gaps 10;

QY 9 DPAPNISEIIOHQYPCBEYEVATEDGYILSVNRIIPGLVOPKKT-----GSRPVVL 62

DB 28 DPELANNTSILIERGWYKAEVHTVTEGYILQMRIFVG-----KTSVTWNGRKPVL 82

QY 63 LQHLVGGASNIISLNPNSLGFILADAGFDVWNGSRGNASRKHKTLISIDQDEFWAFS 122

DB 83 LQHLVGGASNIISLNPNSLGFILADAGFDVWNGSRGNASRKHKTLISIDQDEFWAFS 142

QY 123 YDEMARFELPAPNIFLOKTGQEKIYYGYSGQTMTGFIATFMTPELAQIKMYFALAPI 180

143	Db	WDEMAQYDVDPAMDVHVLAWTGOENLYNMGHSGGTLIMPTHLAKDTCGSPAKKIKRYFALA	202
181	Qy	PIATVKHAKSPGTRKLELLFPDMIMK---LFGKKEFLYQTRFLRQLNLYLQGVILD-QIC	236
203	Db	PIGAVKNIKGLFSYFAHKFSPEFDGWYELFGSKDFLPDNNITKMAAKDICGASEKEAELC	262
237	Qy	SNIMLLGGFNTNNMMGRASVYAHTLACTSVONILHNSQVNSGELRAPDWSGSKL	296
263	Db	DNEFLIGFESDQWNAKRTAIYSSQDPAGTSTQNIHVHMMVRNGRVPAPDFWGGKI-NK	321
297	Qy	EKCNOPTPVRY---RVRDWTVPTAMWTGGQDNLSPEDV-KMLISVNTNLIYHKNI--PE	350
322	Db	KKYQDTPPEYDFGAIKGTKI--HLVSDDDWLDGPTDIDHDFLLKELNPVIAENVNLSK	379
351	Qy	WAHVDIFINGLDAPHRMVNEII	371
380	Db	YNHLDFTWGLNATFGIYDKAI	400

RESULT 7
T33198
hypothetical protein ZK6.7 - Caenorhabditis elegans
C/Species: Caenorhabditis elegans
C/Date: 29-Oct-1999 #sequence_revision 29-Oct-1999 #text_change 21-Jan-2000
C/Accession: T33198
R/Wu. X.
submitted to the EMBL Data Library, May 1998
A/Description: The sequence of C. elegans cosmid ZK6.
A/Reference number: Z21301
A/Accession: T33198
A/Status: preliminary; translated from GB/EMBL/DBJ
A/Molecule type: DNA
A/Residues: 1-403 <WUX>
A/Cross-references: EMBL:AF067942; PIDN:AA017694.1; GSPDB:GNO0023; CESP:ZK6.7
A/Experimental source: strain Bristol N2; clone ZK6
C/Genetics:
A/Gene: CESP:ZK6.7
A/Map position: 5
A/Introns: 29/3; 63/1; 219/3; 365/3
C/Superfamily: triacylglycerol lipase, lingual

```

RESULT 8
T22675  hypothet
C:Specie
C:Date:
C:Access
R:Percy,
submitte
A:Refer
A:Access
A:Status
A:Molecu
A:Residu
A:CROSS-
A:Experi
C:Geneti
A:Gene:
A:Map po
A:Intron
C:Superf

```

```

RESULT 9
H88930
protein R11G11.14 [imported] - Caenorhabditis elegans
C:Species: Caenorhabditis elegans
C:Date: 10-May-2001 #sequence_revision 10-May-2001 #text_change 24-May-2001
C:Accession: H88930
R:anonymous, The C. elegans Sequencing Consortium.
Science 282, 2012-2018, 1998
A:Title: Genome sequence of the nematode C. elegans: a platform for investigating biological processes
A:Reference number: A75000; MUID:R9069613; PMID:9851916
A:Note: see websites genome.wustl.edu/gsc/C.elegans/ and www.sanger.ac.uk/Projects/C_elegans/
A:Note: published errata appeared in Science 283, 35, 1999; Science 283, 2103, 1999; and Science 283, 2104, 1999.
A:Accession: H88930
A:Status: preliminary
A:Molecule type: DNA
A:Residues: 1-405 <STO>
A:Cross-references: GB:chr V; PIDN:AAC69088.1; PID:G2384863; GSPDB:GN00023; CESP:R11G11.1
A:Note: Similar to lipase; R11G11.14

```

C:Genetics:
A:Gene: R11G11.14
A:Map position: 5
C:Superfamily: triacylglycerol lipase, lingual

Query Match 33.7%; Score 699; DB 2; Length 405;
Best Local Similarity 39.5%; Pred. No. 3.1e-54;
Matches 158; Conservative 69; Mismatches 137; Indels 36; Gaps 13;

QY 1 VHMPTKAVDPFAFMNISEIIHQHGYPCPEYEVATEDGYLSVNRIPRGLVQPKKT----55
DB 15 VHVHAKSDPELHMTTPQIIERWGPAMYISVTDDGYILELHRIPEG-----KNTWTP 69
QY 56 -GSRPVLLQGLVGAASNLSNPNNSLGFILADAGFDVVMGNSRGNAWSRKHKTLSID 114
DB 70 NGKQVWFVMOHGLLCASDTWMLPEQSAFIFADAGFDVLMGNRGNTYSMKHKLKAS 129
QY 115 QDEWAFSYDEMARFDLPAVINFILOKTQGEKIYVYVSGQT-TMGFIATSTPE-----L 169
DB 130 HSDFEWSEDEWATYDLPAMDVKVLEVTGQESLYMGHSQGLTWM-----FSLHSDKDGIF 185
QY 170 AQIKMYFALAPIATVGHAKSPGKFLLLPDMIMK-----LFGKKEFLYQTRFLRLQVYL 226
DB 186 AKKIKKFFALAPVGSVKDIKGLFFFAHYFSLFEDGDFVFGAGEFLPNNWAKLAADI 245
QY 227 CGQV-ILDQICSNIMLLGGFTNNMNSRASVVAHTLAGTSVQNLHWSQAVNSGELR 285
DB 246 CGGLKIESDLCDNVCFIAGPESQDNSTRVYVASHDPAQTATONIVHMIQVYRHGGVP 305
QY 286 AFDGSETKMLEKCN-QTPVRYVRDMT-----VPTAMWTGGQDLNPNEDV-KMLSEV 339
DB 306 AYDMSK-ENKNVNFQANPEY---DFTAKGTQIVLYNSDADWLADKDTIYNLLTRL 361
QY 340 TNLHYKX--LPEWAVDFIWLGDAPHMYNEIHLMOQE 377
DB 362 NPAAIQNNYFTDYNHFDVFGRLAPNDIYLPVICTKD 401

RESULT 10
G89074
protein K04A8.5 [imported] - Caenorhabditis elegans
C:Species: Caenorhabditis elegans
C>Date: 10-May-2001 #sequence_revision 10-May-2001 #text_change 24-May-2001
C:Accession: G89074
R:anonymus, The C. elegans Sequencing Consortium.
Science 282, 2012-2018, 1998
A:Title: Genome sequence of the nematode C. elegans: a platform for investigating biology
A:Reference number: A75000; MUID:99069613; PMID:9851916
A>Note: see websites genome.wustl.edu/gsc/C.elegans/ and www.sanger.ac.uk/projects/C.ele
A:Note: published errata appeared in Science 283, 35, 1999; Science 283, 2103, 1999; and
A:Accession: G89074
A>Status: preliminary
A:Molecule type: DNA
A:Residues: 1-411 <STO>
A:Cross-references: GB:chr_V; PIDN:AAC48051.1; PID:g1658354; GSPDB:GN00023; CESP:K04A8.5
C:Genetics:
A:Gene: K04A8.5
A:Map position: 5
C:Superfamily: triacylglycerol lipase, lingual

Query Match 33.3%; Score 691.5; DB 2; Length 411;
Best Local Similarity 38.6%; Pred. No. 1.5e-53;
Matches 150; Conservative 77; Mismatches 143; Indels 19; Gaps 9;

QY 5 TKAVDPEAFPMNISEIIHQHGYPCPEYEVATEDGYLSVNRIPRGLVQPKKT---GSRPVV 61
DB 25 SKSDLEFLYLTPELISWGSVBIYNTTDXGFILELHRIPIYGREVPTSSDVNNSPVI 84
QY 62 LLQHLVGAASNLSNPNNSLGFILADAGFDVVMGNSRGNAWSRKHKTLSIDQD-EFWA 120
DB 85 FLOHGFCLSSFDVANSFQAGVAFADAGFDVLMGNFRGNTYSRKHSVLSNPKDPKFD 144
QY 121 PSYDEMARFDLPAVINFILOKTQGEKIYVYVSGQT-TMGFIATSTPELAQIKMYFALA 180

C:Genetics:
A:Gene: R11G11.14
A:Map position: 5
C:Superfamily: triacylglycerol lipase, lingual

Query Match 31.4%; Score 651.5; DB 2; Length 426;
Best Local Similarity 36.8%; Pred. No. 5.7e-50;
Matches 136; Conservative 82; Mismatches 137; Indels 15; Gaps 8;

QY 7 AYDPEAFPMNISEIIHQHGYPCPEYEVATEDGYLSVNRIP-RLVQPKKTSRPPVILQH 65
DB 20 AIDDECYMTVPPIGKHFGYSEVHLVTTDEYILELHRIPICKQNEKCDRSSKRPVFMQH 79
QY 66 GLVGGASNWSNLPNNSLGFILADAGFDVVMGNSRGNAWSRKHKTLSIDQDEFWAFSYDE 125
DB 80 GLLADGFSWIPNLANSAGVFPADAGFDIWIANSRGTASQKHIGYGPENQKFWNFTWQ 139
QY 126 MARFDLPAVINFILOKTQGEKIYVYVSGQT-TMGFIATSTPELAQIKMYFALAIPATV 185
DB 140 MSEFDLTASVDIVLKTQGEFLYILGHSGQTMIFESRLAENPEFAKIRHFHALAPVATV 199
QY 186 KHAQSP---GTPKFLLLPDMIMKGLFGKKE--LQTRFLRLQVILCGVILDOICSN 239
DB 200 SHIGGLFGLFGKFLTYAEI-----LLGRIPVSPISIPVQKMSYMSRPFMNICITLD 255
QY 240 MLLGGFTNNMNSRASVVAHTLAGTSVQNLHWSQAVNSGELRAPDWGSETKXLEKC 299
DB 256 IGFDIG-NEKMFNQSRVGYLCHTTPATSVKDLQHWIQLVKSQTSKFDYGTD-GNILEY 313
QY 300 NQPTPVRYVRDMTVP-TAMWTGGQDLNPNEDV-KMLSEVNLHYK-NIPWAWHVDPI 357
DB 314 GQTPPEPDLTQINTPTLYLWSRDDILADTDIIRDSILSKMNTIAGSLLEPHYSMDPV 373
QY 358 WGLDAPHRY 367
DB 374 FGTHAFLDY 383

RESULT 12

JT0949
 egg-specific protein - silkworm
 C:Species: Bombyx mori (silkworm)
 C>Date: 31-Mar-1992 #sequence_revision 31-Mar-1992 #text_change 26-Feb-1998
 C:Accession: JT0949; A28527
 R:Sato, Y.; Yamashita, O.
 submitted to JIPID, September 1991
 A:Reference number: JT0949
 A:Accession: JT0949
 A:Molecule type: DNA
 A:Residues: 1-559 <SAT>
 A:Experimental source: larva
 A>Note: this protein is a homotrimer
 R:Indrasith, L.S.; Sasaki, T.; Yamashita, O.
 J. Biol. Chem. 263, 1045-1051, 1988
 A:Title: A unique protease responsible for selective degradation of a yolk protein in B.
 A:Reference number: A28527; PMID:88087166; PMID:3275655
 A:Accession: A28527
 A:Molecule type: protein
 A:Residues: 19-40;133-144, 'T', 146-152;229-248 <IND>
 C:Genetics:
 A:Map position: 19
 C:Keywords: egg yolk; homotrimer
 F:132-133/Cleavage site: Lys-Asn (egg 30.5K cysteine proteinase) #status experimental
 F:228-229/Cleavage site: Arg-Asp (egg 30.5K cysteine proteinase) #status experimental

Query Match 27.6%; Score 574; DB 2; Length 559;
 Best Local Similarity 33.5%; Pred. No. 6.6e-43;
 Matches 130; Conservative 81; Mismatches 149; Indels 28; Gaps 8;

QY 5 TKAVDPEAF-MNISEIIHQGYCEYEVEDGYYLSVNRIPRGLVQPKYTSRPRVLL 63
 DB 180 TQRENNEDFHLNATELLKQGYVEEHTVATDDGYHLTLRIP-PHTRDCKKFPVALL 238
 QY 64 QHGLVGASNIWNLNPNNSIGFTLADAGPDVVMGNSRGNAWRKHKHTLSIDODEFWAFSY 123
 DB 239 MHGLGSADDLLMGPSKSLAYLWLDAGYDVLGNVGRNGKYSRSHVSKHPALNDFWKFEN 298
 QY 124 DENARFDLPVINFILQNTQOEKIYYGVYSGQTTMGFIAPFTMPPELAQIKMYFALAPIA 183
 DB 299 DEIALHDLPAIDHVLDSIQERLHYIGHSQGAATTFALMSEQPSYNEKIVSMHALSPIV 358
 QY 184 TVKHAKS-----PGTKFL-LLPDMWIKGLFKCKEFLVQTRFLRLQVLYLQGVILDO- 234
 DB 359 YMYVRSPLFRMTAPSKFYQYHQVGHGAFEPKHLIET-----FGAACREKL 409
 QY 235 ----ICSNIMLLGGFNTNNMNSRASVYAAHTLAGTSVQNLHLWSQAVNSGELRAFDWG 290
 DB 410 GCRHVCNMLNVISGINVYNQDADIVPVVMAHLPAQTSARVMKQYQGVNASHDFRKYNG 469
 QY 291 SETKYLEKQNPVRYVRVDMVTPTAMWTGGODWLSNPDVQMLLSEVNTLIYHKNIEP 350
 DB 470 AET-NMKVYGASPEPPSDLSKVAPVNLVYSHDAWLAKPKVKELOENTLPNVKQSPVEFE 528
 QY 351 WAH----VDFIWLGLDAPHRMYNEIHLMQ 375
 DB 529 QQHFTDLDFQSKKAPDVTYOKLMENMQ 556

RESULT 13

T39540
 triglyceride lipase-cholesterol esterase - fission yeast (Schizosaccharomyces pombe)
 C:Species: Schizosaccharomyces pombe
 C>Date: 03-Dec-1999 #sequence_revision 03-Dec-1999 #text_change 17-Mar-2000
 C:Accession: T39540
 R:Wood, V.; Rajandream, M.A.; Barrell, B.G.; Beck, A.; Reinhardt, R.
 submitted to the EMBL Data Library, February 1998
 A:Reference number: Z21862
 A:Accession: T39540
 A:Status: preliminary; translated from GB/EMBL/DBJ
 A:Molecule type: DNA

A:Residues: 1-443 <WOO>
 A:Cross-references: EMBL:AL021748; PIDN:CAA16863.1; GSPDB:GN00067; SPDB:SPBCL6A3.12c
 A:Experimental source: strain 972h-; cosmids c16A3
 C:Genetics:
 A:Gene: SPDB:SPBCL6A3.12c
 A:Map position: 2
 A:Introns: 11/2; 152/3; 282/1; 398/3
 C:Superfamily: triacylglycerol lipase, lingual

Query Match 23.7%; Score 491; DB 2; Length 443;
 Best Local Similarity 30.7%; Pred. No. 1.2e-35;
 Matches 117; Conservative 81; Mismatches 135; Indels 48; Gaps 11;

QY 15 NISEIIHQGYCEYEVEDGYYLSVNRIPRGLVQPKYTSRPRVLLQHLGVLGASN 73
 DB 74 NIVEICEAFGYRVEEHLVTRQDNFILLCHRI----THPKSQSKREVYVYCHGLMTNSEL 129
 QY 74 WIS-NLPNNSIGFTLADAGPDVVMGNSRGNAWRKHKHTLSIDODEFWAFSYDEMAFDLP 132
 DB 130 WVAVESERSLPPFVLLIESGYDVMGNNGKYSRKHITYPKDEEFNFSLDMAFDIP 189
 QY 133 AVINFILQKTQOEKIYYGVYSGQTTMGFIAPFTMPPELAQIKMYFALAPIATYVHAKSPG 192
 DB 190 DTVDYILRETGREKLYIGFSQGTQAQMAALSINPLDNKVNIFIGLAPAYAPKGFSNYF 249
 QY 193 TKFLLL--PDMMIKGLFGKKEFLYQTFRLQVLYLQGVILDOICSNIMLLGGFNTNN 250
 DB 250 VDYIVKVPKIMYH-LFGRKCLLPSTVFWQN----ICYPPIFVKIVDVSVKILFNLDLSN 304
 QY 251 MNMSRASVYAAHTLAGTSVQNLHLWSQAVNSGELRAFD-----WGSE-----TK 294
 DB 305 ISLNQKLCGVAHLYSFSSVSVVHMLQIKNCTFQLYDDDMALLAGVGRHYQVPLPPTN 364
 QY 295 NLEKQNPVRYVRVDMVTPTAMWTGGQDWLSNPDVQMLLSEVNTLIYHKNIEPWAHV 354
 DB 365 NI-KC-----FMLILWGKDTLINMEVMTALPPHAKEV---SIAHYEHL 405
 QY 355 DFIWGLDAPHRMYNEIHLMQ 375
 DB 406 DFLAGQDVKEEVPFVVIDALK 426

RESULT 14

T31611
 hypothetical protein Y508A.g - Caenorhabditis elegans
 C:Species: Caenorhabditis elegans
 C>Date: 29-Oct-1999 #sequence_revision 29-Oct-1999 #text_change 29-Oct-1999
 C:Accession: T31611
 R:Steward, C.
 submitted to the EMBL Data Library, September 1999
 A:Reference number: Z21047
 A:Accession: T31611
 A:Status: preliminary; translated from GB/EMBL/DBJ
 A:Molecule type: DNA
 A:Residues: 1-1585 <WIL>
 A:Cross-references: EMBL:AL117200; NID:el549770; PIDN:CAB55050.1; CESP:Y508A.g
 A:Experimental source: clone Y508A
 C:Genetics:
 A:Gene: CESP:Y508A.g
 A:Introns: 25/3; 60/1; 133/2; 217/3; 270/3; 337/2; 400/1; 746/2

Query Match 23.0%; Score 478.5; DB 2; Length 1585;
 Best Local Similarity 30.6%; Pred. No. 9e-34;
 Matches 118; Conservative 67; Mismatches 172; Indels 29; Gaps 9;

QY 11 EAFNISEIIHQGYCEYEVEDGYYLSVNRIPRGLVQPKYTSRPRVLLQHLGVL 68
 DB 20 EETLDAADTISHYGYVVEKHVTTDDGYTVQLQRIPVGRDDRSILGCKEPVVFVHGLF 79
 QY 69 CGASNWISNLPNNSIGFTLADAGPDVVMGNSRGNAWRKHKHTLSIDODEFWAFSYDEMA 128
 DB 80 GSSYHFLNLPSSQAAYIFADAGPDVVMGLNGIRGTGYGLNHTSPSTNGVNFNFSLEYHSH 139

Search completed: February 19, 2004, 14:23:08
Job time : 15.4317 secs

```

Qy 129 FDLPAVINFIQKQEQKIYVYGSQGTMTGFIATSTMPEL-----AKIKWYFALAPIAT 184
Db 140 YDLRQQLYYVLDYTRHESLFVGHSGCTA---VMFARLAADVTWQSKIYVFFALGPTAG 196
Qy 185 VKHAKSPGTFKLLLPDMIKGLPKGKFLYQTRP-----LRQLVYLCGVILDO 234
Db 197 FMKPLMP---FTLEENYLAQI---QFALDGRGFLPVEIPRAIASKPADFCSSKFPTF 250
Qy 235 ICNIMILLGGFNT--NNMNSRASVYAHTLAGTSVONILHWSQAVNSGELRAFDWGSSET 293
Db 251 LCSAGPKVAAGIETLQGVNDSRIPILLSHPSPATSITLNLHWNQIPKYHELRLDLGT-A 309
Qy 294 KNLEKCNQPTPVVRVDMVTPTAMWTGGQDMLSNPDEDVOLLSEV---THLIYHKNIPEW 351
Db 310 RNLIAQGQKAPRIEIGNIATQILYFSKDDQITDEVDVREIIMKMGPGGLIESYDLDFH 369
Qy 352 AHVDPIWGLDAPHRMNNEIHLMOOE 377
Db 370 THDFILGLRATDEVYKPIVYRIYKE 395

RESULT 15
T41053
triglyceride lipase-cholesterol esterase - fission yeast (Schizosaccharomyces pombe)
C:Species: Schizosaccharomyces pombe
C:Date: 03-Dec-1999 #sequence_revision 03-Dec-1999 #text_change 31-Mar-2003
C:Accession: T41053
R:Halbert, H.; Duesterhoeft, A.; Wood, V.; Rajandream, M.A.; Barrell, B.G.
submitted to the EMBL Data Library, August 1998
A:Reference number: Z21967
A:Accession: T41053
A:Status: Preliminary; translated from GB/EMBL/DBJ
A:Molecule type: DNA
A:Residues: 1-467 <HIL>
A:Cross-references: EMBL:AL031324; PIDN:CAA0447.1; GSPDB:GN00068; SPDB:SPCC1672.09
A:Experimental source: strain 972h-; cosmid c1672
C:Genetics:
A:Gene: SPDB:SPCC1672.09
A:Map position: 3
A:Introns: 12/2; 164/3; 294/1
C:Superfamily: triacylglycerol lipase, lingual

Query Match 22.9%; Score 475; DB 2; Length 467;
Best Local Similarity 31.2%; Pred. No. 3.4e-34;
Matches 125; Conservative 68; Mismatches 158; Indels 50; Gaps 12;

Qy 6 KAVDPEAFMNISEIIHQGYCEYEYVATEDGYILSVNRIPR-----GLVQPKYTSRSP 59
Db 73 KAIHDA--DIRECKISGYTVEDHLVTRTDDYILCHIRISKOSPGRIGSPHKKL---P 127
Qy 60 VVLLQHLGIVGASWISNL--PNNSLGFILADAGFDVWNGSRGNASRKHKTLSIDQDEF 118
Db 128 VVYCHGLLMNSEVVCNVDPNCLVFDLVNKGVDVWLNNGNNGKYSRQHLRFDSTDKF 187
Qy 119 WAFSYDEMAREDLPAVINFIQKQEQKIYVYGSQGTMTGFIATSTMPELAKIKWYFA 178
Db 188 WDFSIDDPAQYDIPDTIDYILKTSQTKLTLYIGFSQGTAFAPASLSIHPILNDKINSLIA 247
Qy 179 LAPATVNHAKSPGTFKLLPDMWIKG-----LFGKGFYOTRFLRQLVYLCQVI 231
Db 248 LAP-----AISPGLHNVRVDAFVKARPSILFFLGKKSILPSAGFWQSFL-----APKF 297
Qy 232 LDQICSNIMILLGGFNTNNMNSRASVYAHTLAGTSVONILHWSQAVNSGELRAFD--- 288
Db 298 FDAVLAYCLSQLEFNWSCQNISSYQRLVSFAHLYSYTSVKCLVHWFQIMRSAEFMYDNDQ 357
Qy 289 WGSE--TKNLEKCNQPTPVVRVDMVTPTAMWTGGQDMLSNPDEDVOLLSEVTLNLIYHK 346
Db 358 LGHDYFLKYKAAKEPT-----NNIRTPYLIWGGSDSL---VDIQAMNLPAEVEHV 408
Qy 347 NIPWAHVDFIWG-----LDAPHRMNEIHLMOOEETN 380
Db 409 KVDSEHLDMIWADTVKDYVPPVPLRLRLDIHPPHEEND 449

```


GenCore version 5.1.6
Copyright (c) 1993 - 2004 Compugen Ltd.

OM protein - protein search, using sw model

Run on: February 19, 2004, 14:17:14 ; Search time 27.8229 Seconds
(without alignments)
3617.186 Million cell updates/sec

Title: US-10-042-431-49

Perfect score: 2076

Sequence: 1 VHMPTKAVDEAFNISEII.....IHLMQEETNLQGRCEAVL 390

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 830525 seqs, 258052604 residues

Total number of hits satisfying chosen parameters: 830525

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database :

SPTREMBL 23:*

- 1: sp_archaea:*
- 2: sp_bacteria:*
- 3: sp_fungi:*
- 4: sp_human:*
- 5: sp_invertebrate:*
- 6: sp_mammal:*
- 7: sp_mhc:*
- 8: sp_organelle:*
- 9: sp_phase:*
- 10: sp_plant:*
- 11: sp_rodent:*
- 12: sp_virus:*
- 13: sp_vertebrate:*
- 14: sp_unclassified:*
- 15: sp_virus:*
- 16: sp_bacteriap:*
- 17: sp_archaeap:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	1991	95.9	374	4 Q96LG2	Q96LG2 homo sapien
2	1848.5	89.0	422	11 Q8K2A6	Q8K2A6 mus musculus
3	1844.5	88.8	422	11 Q8BJ28	Q8BJ28 mus musculus
4	1219.5	58.7	397	11 Q8C3G7	Q8C3G7 mus musculus
5	1130	54.4	395	11 Q8D798	Q8D798 mus musculus
6	1130	54.4	395	11 Q9D6G8	Q9D6G8 mus musculus
7	1130	54.4	395	11 Q9D6P3	Q9D6P3 mus musculus
8	1130	54.4	395	11 Q9D6X0	Q9D6X0 mus musculus
9	1129	54.4	395	11 Q9C8P7	Q9C8P7 mus musculus
10	1129	54.4	395	11 Q9C8P8	Q9C8P8 mus musculus
11	1127	54.3	395	11 Q9D6L9	Q9D6L9 mus musculus
12	1125	54.2	395	11 Q9D6T5	Q9D6T5 mus musculus
13	1124	54.1	395	11 Q9D767	Q9D767 mus musculus
14	1123	54.1	395	11 Q9D7C5	Q9D7C5 mus musculus
15	1122	54.0	395	11 Q9D6L1	Q9D6L1 mus musculus
16	1122	54.0	395	11 Q9D766	Q9D766 mus musculus

ALIGNMENTS

RESULT 1

Q96LG2 ID Q96LG2 PRELIMINARY; PRT; 374 AA.

AC Q96LG2; 19, Created)
DT 01-DEC-2001 (TREMBLrel. 19, Last sequence update)
DT 01-DEC-2001 (TREMBLrel. 19, Last sequence update)
DT 01-MAR-2003 (TREMBLrel. 23, Last annotation update)
DE BA30415.1 (Novel lipase) (Fragment).
GN BA30415.1.
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
OX NCBI_TaxID=9606;
RN [1]
RP SEQUENCE FROM N.A.
RA Bray-Allen S.;
RL Submitted (AUG-2001) to the EMBL/GenBank/DBJ databases.
DR EMBL; AL353113; CAC78754.1; -
DR InterPro; IPR000073; A/b hydrolase.
DR InterPro; IPR006693; abhydro_lipase.
DR InterPro; IPR000734; Lipase.
DR InterPro; IPR000379; Ser_estr_site.
DR Pfam; PF00561; abhydro_lipase; 1.
DR Pfam; PF04083; abhydro_lipase; 1.
DR PROSITE; PS00120; LIPASE_SER; 1.
FT NON_TER 1
SQ SEQUENCE 374 AA; 42371 MW; 5E7220A889437337 CRC64;

Query Match 95.9%; Score 1991; DB 4; Length 374;
Best Local Similarity 100.0%; Pred. No. 1.3e-175;
Matches 374; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 17 SEIIHQGPCEYEYVATEDGVILSVNRPGLVQPKTGRPVVLLQHLGVGGASNNIS 76
Db 1 SEIIHQGPCEYEYVATEDGVILSVNRPGLVQPKTGRPVVLLQHLGVGGASNNIS 60
QY 77 NLPNNSLGFILADAGFDVMMGNMGRNNAWRKHTLSIDQDEFWAFSYDEMAFDPILPAVIN 136
Db 61 NLPNNSLGFILADAGFDVMMGNMGRNNAWRKHTLSIDQDEFWAFSYDEMAFDPILPAVIN 120

QY 137 FILQKTGQEKIYYVGSQGTGFIAPSTMPPELAOKIKMYFALAPIATVKAHAKSPGTRFL 196
 Db 121 FILQKTGQEKIYYVGSQGTGFIAPSTMPPELAOKIKMYFALAPIATVKAHAKSPGTRFL 180
 QY 197 LLPDMKIKGLFGKKEFLYQTRFLRQLVYLCOGQVILDOICSNIMLLGGFNTNNMGRSA 256
 Db 181 LLPDMKIKGLFGKKEFLYQTRFLRQLVYLCOGQVILDOICSNIMLLGGFNTNNMGRSA 240
 QY 257 SVTAHTLAGTSVONILHWSQAVNSGELRAFDWGSSTKLEKCNQPTPVRYVRDVTPT 316
 Db 241 SVTAHTLAGTSVONILHWSQAVNSGELRAFDWGSSTKLEKCNQPTPVRYVRDVTPT 300
 QY 317 AMWTGGQDWLSPNEDVKMLSEVTNLIYHKNIPWAHVDVFWGLDAPHMYNEIHLMOQ 376
 Db 301 AMWTGGQDWLSPNEDVKMLSEVTNLIYHKNIPWAHVDVFWGLDAPHMYNEIHLMOQ 360
 QY 377 EETNLSQGRCEAVL 390
 Db 361 EETNLSQGRCEAVL 374

RESULT 2

Q8K2A6 PRELIMINARY; PRT; 422 AA.
 AC Q8K2A6
 DT 01-OCT-2002 (TREMELrel. 22, Created)
 DT 01-OCT-2002 (TREMELrel. 22, Last sequence update)
 DT 01-MAR-2003 (TREMELrel. 23, Last annotation update)
 DE RIKEN CDNA 4632427C23 gene (BA30415.1).
 OS Mus musculus (Mouse).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Mammalia; Euthera; Rodentia; Sciurognathi; Muridae; Mus.
 OX NCBI_TaxID=10090;
 RN [1]
 RP SEQUENCE FROM N.A.
 RA Strausberg R.;
 RL Submitted (JUN-2002) to the EMBL/GenBank/DBJ databases.
 RN [2]
 RP SEQUENCE FROM N.A.
 RC STRAIN=C57BL/6J; TISSUE=Mammary gland, and Vagina;
 RC MEDLINE=22354683; PubMed=12466851;
 EX The FANTOM Consortium,
 RA the RIKEN Genome Exploration Research Group Phase I & II Team;
 RT "Analysis of the mouse transcriptome based on functional annotation of
 RT 60,770 full-length cDNAs.";
 RL Nature 420:563-573(2002).
 DR EMBL; BC031933; AAH31933.1; -
 DR EMBL; AK037031; BAC29899.1; -
 DR EMBL; AK085719; BAC39517.1; -
 DR InterPro; IPR000073; A/b hydrolase.
 DR InterPro; IPR006693; abhydro_lipase.
 DR InterPro; IPR000734; Lipase.
 DR InterPro; IPR000379; Ser esterase.
 DR Pfam; PF00561; abhydro_lipase; 1.
 DR Pfam; PF04083; abhydro_lipase; 1.
 DR PROSITE; PS00120; LIPASE_SER; 1.
 SQ SEQUENCE 422 AA; 48253 MW; C6BFEC140963E051 CRC64;

Query Match 89.8%; Score 1848.5; DB 11; Length 422;
 Best Local Similarity 88.4%; Pred. No. 2.4e-162;
 Matches 344; Conservative 20; Mismatches 24; Indels 1; Gaps 1;
 QY 2 HMPKAVDPEAFMNIISIIHQGYPCBEYEVATEDGYILSVNRIPRGLVOPKKTGSRPVV 61
 Db 35 HLPTKAADPEAFMNVSEIIKHGYPSEYEVATEDGYILSVNRIPRGQTLKKEGSRPVV 94
 QY 62 LLOHGLVGASNWSIPLNPNNSLGFILADAGFDVVMGNSRGNSRKHKTLSIDQDEFWAF 121
 Db 95 LLOHGLVGASNWSIPLNPNNSLGFILADAGFDVVMGNSRGNSRKHKTLSIDQDEFWAF 154
 QY 122 SYDEMARFDPVAVINFILQKTGQEKIYYVGSQGTGFIAPSTMPPELAOKIKMYFALAP 181

Db 155 SYDEMARFDPVAVINFILQKTGQEKIYYVGSQGTGFIAPSTMPPELAOKIKMYFALAP 214
 QY 182 IATVKAHAKSPGTRFLRQLVYLCOGQVILDOICSNIMLLGGFNTNNMGRSA 241
 Db 215 IATVKAHAKSPGTRFLRQLVYLCOGQVILDOICSNIMLLGGFNTNNMGRSA 274
 QY 242 LLOGFNTNNMGRSAVYAHTLAGTSVONILHWSQAVNSGELRAFDWGSSTKLEKCNQ 301
 Db 275 LLOGFNTNNMGRSAVYAHTLAGTSVONILHWSQAVNSGELRAFDWGSSTKLEKCNQ 334
 QY 302 PTPVRYVRDVTPTAMWTGGQDWLSPNEDVKMLSEVTNLIYHKNIPWAHVDVFWGLD 361
 Db 335 PTPVRYVRDVTPTAMWTGGQDWLSPNEDVKMLSEVTNLIYHKNIPWAHVDVFWGLD 394
 QY 362 APHRMYNEIHLMOQEEETNLSQGRCEAVL 390
 Db 395 APQRVYNEIHLMAKQ-EPNLPQGTCTRVKL 422

RESULT 3

Q8BU28 PRELIMINARY; PRT; 422 AA.
 AC Q8BU28
 DT 01-MAR-2003 (TREMELrel. 23, Created)
 DT 01-MAR-2003 (TREMELrel. 23, Last sequence update)
 DT 01-MAR-2003 (TREMELrel. 23, Last annotation update)
 DE BA30415.1
 OS Mus musculus (Mouse).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Mammalia; Euthera; Rodentia; Sciurognathi; Muridae; Mus.
 OX NCBI_TaxID=10090;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN=C57BL/6J; TISSUE=Skin;
 RC MEDLINE=22354683; PubMed=12466851;
 EX The FANTOM Consortium,
 RA the RIKEN Genome Exploration Research Group Phase I & II Team;
 RT "Analysis of the mouse transcriptome based on functional annotation of
 RT 60,770 full-length cDNAs.";
 RL Nature 420:563-573(2002).
 DR EMBL; AK037214; BAC29757.1; -
 SQ SEQUENCE 422 AA; 48267 MW; 5F2522D0E27050C8 CRC64;

Query Match 88.8%; Score 1844.5; DB 11; Length 422;
 Best Local Similarity 88.2%; Pred. No. 5.7e-162;
 Matches 343; Conservative 20; Mismatches 25; Indels 1; Gaps 1;
 QY 2 HMPKAVDPEAFMNIISIIHQGYPCBEYEVATEDGYILSVNRIPRGLVOPKKTGSRPVV 61
 Db 35 HLPTKAADPEAFMNVSEIIKHGYPSEYEVATEDGYILSVNRIPRGQTLKKEGSRPVV 94
 QY 62 LLOHGLVGASNWSIPLNPNNSLGFILADAGFDVVMGNSRGNSRKHKTLSIDQDEFWAF 121
 Db 95 LLOHGLVGASNWSIPLNPNNSLGFILADAGFDVVMGNSRGNSRKHKTLSIDQDEFWAF 154
 QY 122 SYDEMARFDPVAVINFILQKTGQEKIYYVGSQGTGFIAPSTMPPELAOKIKMYFALAP 181
 Db 155 SYDEMARFDPVAVINFILQKTGQEKIYYVGSQGTGFIAPSTMPPELAOKIKMYFALAP 214
 QY 182 IATVKAHAKSPGTRFLRQLVYLCOGQVILDOICSNIMLLGGFNTNNMGRSA 241
 Db 215 IATVKAHAKSPGTRFLRQLVYLCOGQVILDOICSNIMLLGGFNTNNMGRSA 274
 QY 242 LLOGFNTNNMGRSAVYAHTLAGTSVONILHWSQAVNSGELRAFDWGSSTKLEKCNQ 301
 Db 275 LLOGFNTNNMGRSAVYAHTLAGTSVONILHWSQAVNSGELRAFDWGSSTKLEKCNQ 334
 QY 302 PTPVRYVRDVTPTAMWTGGQDWLSPNEDVKMLSEVTNLIYHKNIPWAHVDVFWGLD 361
 Db 335 PTPVRYVRDVTPTAMWTGGQDWLSPNEDVKMLSEVTNLIYHKNIPWAHVDVFWGLD 394
 QY 362 APHRMYNEIHLMOQEEETNLSQGRCEAVL 390

```
Db 395 AQRVYNEIHLKQ-EPNLPQGTCTVKL 422
RESULT 4
ID Q8C2G7 PRELIMINARY; PRT; 397 AA.
AC Q8C2G7;
DT 01-MAR-2003 (TReMBLrel. 23, Created)
DT 01-MAR-2003 (TReMBLrel. 23, Last sequence update)
DT 01-MAR-2003 (TReMBLrel. 23, Last annotation update)
DE Lysosomal acid lipase 1.
OS Mus musculus (Mouse).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
OX NCBI_TaxID=10090;
[1]
RN SEQUENCE FROM N.A.
RP STRAIN=NOD; TISSUE=Thymus;
RC MEDLINE=22354683; PubMed=12466851;
RA The FANTOM Consortium,
RA the RIKEN Genome Exploration Research Group Phase I & II Team;
RT "Analysis of the mouse transcriptome based on functional annotation of
RT 60,770 full-length cDNAs.";
RL Nature 420:563-573 (2002).
DR EMBL; AK888659; BAC0484.1; -.
SQ SEQUENCE 397 AA; 45325 MW; 866B123A1A058134 CRC64;
Query Match 58.4%; Score 1219.5; DB 11; Length 397;
Best Local Similarity 58.4%; Pred. No. 3.5e-104;
Matches 220; Conservative 66; Mismatches 88; Indels 3; Gaps 1;
QY 3 MPT---KAVDPKPNVISEIIHQGPCEYEVEVATEDGYILSVNRIIPRGLVQPKTGRSP 59
19 VPTGTVAVDPEVNNVTEIIMRWGYPGEHSVLTGDTYILSHIRIPGRKNHFGKGRPP 78
Db 60 VVLLQHLGVLGASNIWNLNPNLSGLFILADAGFDVWNGSRGNASRKHKTLSDIDQDFW 119
79 VVYLQHLGLADSSNWTNIDNSSLGFLADAGFDVWNGSRGNTWSLKHKTLVSQDFW 138
QY 120 AFSVDEMARPLDAVINFILOKTOEKIYVYGSGCTMTGFIAPSTPELAOKIKMYFAL 179
139 AFSVDEMAKYLPAISINYLKNTQEQIYVYGSGCTTIGFIAPSQMPELAKIKMYFLV 198
Db 180 APIATVHKASFGTKFLLPDMTKGLFGKKEFLYQTRFLRQLVYLCOGVILDOICSN 239
199 APVLSLNFASGLQLGLPLDPLKDMFGQKFLPQSNMLKWLSHVCTHWIKELCANV 258
QY 240 MLLGGFNNNMSRASVYAAHTLAGTSVQNIHLWQVNSGELRAPDWSGSETNLEK 299
259 FFLCGFNEKLNLSRVDVYTHCPAGTSVQNMHLHWGQVFKYKLOAPDWSGSEKRYFY 318
QY 300 NQPTPVRYRDMTPTAMWTGGODWLSNPEDVKMLLSEVNTLIYHKNIPEWAHVDFI 359
319 NQSPFPNINKNRMLPTALNSGGDWLADINDIITLTQPKLYHKNIPEWHDLDPI 378
QY 360 LDAPHMYNEIHLMOQ 376
Db 379 LDAPKLYDEIISLMKX 395
RESULT 5
Q9D798 PRELIMINARY; PRT; 395 AA.
ID Q9D798;
AC Q9D798;
DT 01-JUN-2001 (TReMBLrel. 17, Created)
DT 01-JUN-2001 (TReMBLrel. 17, Last sequence update)
DT 01-MAR-2003 (TReMBLrel. 23, Last annotation update)
DE 2310051B2IRik protein.
GN 2310051B2IRIK.
OS Mus musculus (Mouse).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
OX NCBI_TaxID=10090;
[1]
RN SEQUENCE FROM N.A.
RP STRAIN=CS7BL/6J; TISSUE=Tongue;
RC MEDLINE=21085660; PubMed=11217851;
RA Kawai J., Shinagawa A., Shibata K., Yoshino M., Itoh M., Ishii Y.,
RA Arakawa T., Hara A., Fukunishi Y., Konno H., Adachi J., Fukuda S.,
RA Aizawa K., Izawa M., Nishi K., Kiyosawa H., Kondo S., Yamanaka I.,
RA Saito T., Okazaki Y., Gojobori T., Bono H., Kasukawa T., Saito R.,
RA Kadota K., Matsuda H.A., Ashburner M., Batalov S., Casavant T.,
RA Fleischmann W., Gaasterland T., Gissi C., King B., Kochiwa H.,
RA Kuehl P., Lewis S., Matsumoto Y., Nikaudo I., Pesole G., Quackenbush J.,
RA Schriml L.M., Staubli F., Suzuki R., Tomita M., Wagner L., Washio T.,
RA Sakai K., Okido T., Furuno M., Aono H., Baldarelli R., Barsh G.,
RA Blake J., Boffelli D., Bojunga N., Carninci P., de Bonaudo M.F.,
RA Brownstein M.J., Bult C., Fletcher C., Fujita M., Gariboldi M.,
RA Gustincich S., Hill D., Hofmann M., Hume D.A., Kamiya M., Lee N.H.,
RA Lyons P., Marchionni L., Mashima J., Mazzarelli J., Mombaerts P.,
RA Nordone P., Ring B., Ringwald M., Rodriguez I., Sakamoto N.,
RA Sasaki H., Sato K., Schoenbach C., Seya T., Shibata Y., Storch K.-F.,
RA Suzuki H., Toyooka K., Wang K.H., Weitz C., Whittaker C., Wilming L.,
RA Wyshaw-Boris A., Yoshida K., Hasegawa Y., Kawaji H., Kohsaki S.,
RA Hayashizaki Y.;
RT "Functional annotation of a full-length mouse cDNA collection.";
RL Nature 409:685-690 (2001).
DR EMBL; AK009431; BAB26283.1; -.
DR HSSP; P07098; IHLG.
DR MGD; MGI:1914967; 2310051B2IRik.
DR InterPro; IPR000073; A/b hydrolase.
DR InterPro; IPR000734; Lipase.
DR InterPro; IPR000379; Ser esterase site.
DR Pfam; PF00561; abhydrolase_1.
DR PROSITE; PS00120; LIPASE SER; 1.
SQ SEQUENCE 395 AA; 44603 MW; D3FD8B6FEA671E3E CRC64;
Query Match 54.4%; Score 1130; DB 11; Length 395;
Best Local Similarity 56.2%; Pred. No. 6.7e-96;
Matches 209; Conservative 64; Mismatches 93; Indels 6; Gaps 2;
QY 9 DPEAFNMTSEIIHQGPCEYEVEVATEDGYILSVNRIIPRGLVQPKTGRSPVLLQHLG 68
27 NPEANNVSMQITVYWGYPSEYEVVTEGYILGVTRIPYKKNSENIGKRPVAYLQHLI 86
QY 69 GGASNIWNLNPNLSGLFILADAGFDVWNGSRGNASRKHKTLSDIDQDFWAFSDENAR 128
87 ASATNWTNLPNLSLAFILADAGYDVLGNSRGNTWSKNVYSPDSVEFWAFSDENAK 146
129 FDLPAVINFILOKTOEKIYVYGSGCTMTGFIAPSTPELAOKIKMYFALAPATYKHA 188
147 YDLPATIDFIVOKTQEKIHYVGHSGTIGFIAPSTPNALAKIKRFPALAPATVYKT 206
QY 189 KSPGTFKLLPDMTKGLFGKKEFLYQTRFLRQLVYLCOGVILDOICSNIMLLGG 245
207 ESPFKKISLIPKFLKLVFGNKFMPHNYLDQFLGTEV---CSRELLDLCSNALFICG 263
QY 246 FNTNNMNSRASVYAAHTLAGTSVQNIHLWQVNSGELRAPDWSGSETNLEKQNPV 305
264 FDKNLNLSRDLVGHENPAGTSTQDLFWAQLAKSGKLOAYNNGSPLOQMLHYNKTPT 323
QY 306 RYKVRDMTPTAMWTGGODWLSNPEDVKMLLSEVNTLIYHKNIPEWAHVDFI 365
324 YDVSNTVPIAVNNGGHDILADPDQVAKLLPKFLNLYHKELLYHNDLFIWADAPQE 383
QY 366 MYNEIHLMOQ 377
Db 384 VYNEIVTMAED 395
RESULT 6
Q9D6Q6 PRELIMINARY; PRT; 395 AA.
ID Q9D6Q6;
AC Q9D6Q6;
DT 01-JUN-2001 (TReMBLrel. 17, Created)
DT 01-JUN-2001 (TReMBLrel. 17, Last sequence update)
```

RESULT 7	Q9DP63	PRELIMINARY;	PRT;	395 AA.
ID	Q9DP63			
AC	Q9DP63			
DT	01-JUN-2001 (TRENburel. 17, Created)			
DT	01-JUN-2001 (TRENburel. 17, Last sequence update)			
DT	01-WAR-2003 (TRENburel. 23, Last annotation update)			
DE	Adult male tongue cDNA, RIKEN full-length enriched library,			
DE	clone:2310069P19, full insert sequence.			
GN	2310051B2IRIK			
OS	Mus musculus (Mouse)			
OC	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;			
OC	Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.			
OX	NCBI_TaxId=10090;			
RN	[1]			
RP	SEQUENCE FROM N.A.			
RC	STRAIN=C57BL/6J; TISSUE=Tongue;			
RC	MEDLINE=257BL/6J; PubMed=1121781;			
RX	Kawai J., Shinagawa A., Shibata K., Yoshino M., Itoh M., Ishi Y.,			
RA	Arakawa T., Hara A., Fukunishi Y., Konno H., Adachi J., Fukuda S.,			
RA	Aizawa K., Izawa M., Nishi K., Kiyosawa H., Kondo S., Yamataka I.,			
RA	Saito T., Okazaki Y., Gojobori T., Bono H., Kasukawa T., Saito R.,			
RA	Kadota K., Matsuda H.A., Ashburner M., Batalov S., Caavant T.,			
RA	Pletschmann W., Gaasterland T., Gissi C., King B., Kochiwa H.,			
RA	Kuehl P., Lewis S., Matsuo Y., Nikaido I., Pesole G., Quackenbush J.,			
RA	Schriml L.M., Straubi P., Suzuki R., Tomita M., Wagner L., Washio T.,			
RA	Sakai K., Okido T., Furuno M., Aono H., Baldarelli R., Barsh G.,			
RA	Blake J., Boffelli D., Bojunga N., Carninci P., de Bonaldi M.P.,			
RA	Brownstein M.J., Bult C., Fletcher C., Fujita M., Gariboldi M.,			
RA	Gustingich S., Hill D., Hofmann M., Hume D.A., Kamiya M., Lee N.H.,			
RA	Lyons P., Marchionni L., Mashima J., Mazzarelli J., Mombaerts P.,			
RA	Nordone P., Ring B., Ringwald M., Rodriguez I., Sakamoto N.,			
RA	Sasaki H., Sato K., Schoenbach C., Seya T., Shibata Y., Storch K.-F.,			
RA	Suzuki H., Toyoko-Oka K., Wang K.H., Weitz C., Whittaker C., Wilming L.,			
RA	Wyshaw-Boris A., Yoshida K., Hasegawa Y., Kawaji H., Kohtsuki S.,			
RA	Hayashizaki Y.,			
RT	Functional annotation of a full-length mouse cDNA collection.;			
RL	Nature 409:685-690(2001).			
RL	EMBL; AK010139; BAB26725.1; -			
DR	HSSP; P07098; 1HLG.			
DR	MGD; MGI:1914967; 2310051B2IRIK.			
DR	InterPro; IPR000073; A/b_hydrolase.			
DR	InterPro; IPR000734; Lipase.			
DR	InterPro; IPR000379; Ser_estere_site.			
DR	Pfam; PF00561; abhydrolase; 1.			
DR	DR PROSITE; PS00120; LIPASE_SER; 1.			
DR	SEQUENCE 395 AA; 44651 MW; E48A73CCCBFD359D CRC64;			
Qy	Query Match	54.4%;	Score 1130;	DB 11; Length 395;
Qy	Best Local Similarity	56.5%;	Pred. No. 6.7e-96;	
Qy	Matches	210;	Conservative	63; Mismatches 93; Indels 6; Gaps 2;
Qy	9	DPFAFWNISETIQHGTCPEEYEVATDDGVLISVNIPIRGLVQPKTSRPPVLLQHLV	68	
Db	27	NPEANNVMSQMTIWTGYPSEYEVVTDGVLGYIPYIPYKKNSENIGKRPVAYIQHGLI	86	
Qy	69	GCASWISNLPNLSLGFILADAGDGVWNGSRGNAMSRKHKLISIDQDFWAFSPYDEMAR	128	
Db	87	ASATWITNLPNLSLAFILADAGDGVWLGNSRGNTWSKNTVYSPDSVEFWAFSDEMAK	146	
Qy	129	FDLPAVINFIQKTQGEKIYVGVSGQTWVGFIATFTMPBELAQIKYKVPALAPIATVRA	188	
Db	147	YDLPATIDFIVQKTQGEKIHYVGVSGQTWVGFIATFTMPBELAQIKYKVPALAPIATVRA	206	
Qy	189	KSPGTGKFLLLPDMMLKGLFGKKFPL---YQTRFLRQVLVYLCQVILDOICSNIMLLGG	245	
Db	207	ESPPFKISLIPKFLKLVIFGNMFMFHYLDQFLGTEV---CSRLLDLLCSNLFICG	263	
Qy	246	FTNNNNMRSASVYAHTLAGTSVQNTILHWSQAVNGSELRAFPWGSETKNSKCNQPTPV	305	
Db	264	FDKCNLNSRFDVYLGNHNPAGTSTQDLFWHAQLAKSKGLQAVNWGSLQNLMLHYNQKTPP	323	


```

DR EMBL; AK010125; BAB2671.1; -.
DR HSP; P07098; 1HLG
DR MGD; MGI:1914967; 2310051B21Rik.
DR InterPro; IPR000073; A/B_hydrolase.
DR InterPro; IPR000379; Ser_estra_site.
DR Pfam; PF00561; abhydrolase; 1.
DR PROSITE; PS00120; LIPASE SER; 1.
SQ SEQUENCE 395 AA; 44637 MW; D3F96B65EA671B34 CRC64;

Query Match 54.4%; Score 1129; DB 11; Length 395;
Best Local Similarity 56.2%; Pred. No. 8.3e-96;
Matches 209; Conservative 64; Mismatches 93; Indels 6; Gaps 2;

QY 9 DPEAFMISIIHQGYPCBEYEVATEDGYILSVNRIPLRGVOPKKTGSRPVVLLQHLV 68
DB :|||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
27 NPEANMVSQWITWGYPSSEYEVATEDGYILGVYRIPYKKNSENIKRPVAYLQHLI 86
QY 69 GGASNWTISLNNLSGLIFADAGFDVWGNRGNVSRKHTLSIDQDFWAFSYDEMAR 128
DB :|||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
87 ASATNWTITLNNLSLAFILADAGYDVLGNSRGTWSRKNVYSPDSVEFWAFSDEMAK 146
QY 129 FDLPAVINFLIQKTQGEKIYVGVSGQTMTGFTAFSTMPBELAKIKMYFALAPIATVCHA 188
DB :|||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
147 YDLPATIDFIVOKTGQEKIHVGHSGQTITGFTAFSTNPALAKKIKFPYALAPVATVKYT 206
QY 189 KSPGTFKLLPDMMIKGLFGKKEPL---YOTRFLRQLVYVLCQVILDOICSNIMLLGG 245
DB :|||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
207 ESPFKKISLIPKFLKIFGKMFMPHNYLDQFLGTEV---CSRLLDLCLSNALFIFCG 263
QY 246 FNTNNMNSRASVVAHTLAGTSVQNTILHWSQVNSGELRAFDPWGSSTKLEKCNQPTPV 305
DB :|||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
264 FDKQLNVSRFVYLGHNPGTSTQDLFHWALAKSGKQLQAYNGSPQLQMLHYNQKTPP 323
QY 306 RYVRDMVTPTAMTGGQDMLSNPDEVKMLLSEVTNLIYHNKIPENAHVDFIWLGDAPHR 365
DB :|||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
324 YDVSAMTPTVIAVWNGGHILADPDQVAMLLPKLPLNLLYHKEILPYNHLDFIWMADAPQE 383
QY 366 MYNEIHLMOQE 377
DB :|||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
384 VINEIVTMMAED 395

RESULT 11
Q9D6L9
ID Q9D6L9 PRELIMINARY; PRT; 395 AA.
AC Q9D6L9
DT 01-JUN-2001 (TrEMBLrel. 17, Created)
DT 01-JUN-2001 (TrEMBLrel. 17, Last sequence update)
DE Adult male tongue cDNA, RIKEN full-length enriched library,
DE clone:231007613, full insert sequence.
GN 2310051B21Rik.
OS Mus musculus (Mouse).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
OX NCBI_TaxID=10090;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=C57BL/6J; TISSU=Tongue;
EX MEDLINE=21085660; PubMed=11217851;
RA Kawai J., Shinagawa A., Shibata K., Yoshino M., Itoh M., Ishii Y.,
RA Aizawa K., Iwawa M., Nishi K., Kiyosawa H., Konno H., Adachi J., Fukuda S.,
RA Kadota K., Matsuda H.A., Ashburner M., Batalov S., Casavant T.,
RA Fleischmann W., Gaasterland T., Gissi C., King B., Kochiwa H.,
RA Kuehl P., Lewis S., Matsuo Y., Nikaido I., Pesole G., Quackenbush J.,
RA Schriml L.M., Stauber I. F., Suzuki R., Tomita M., Wagner L., Washio T.,
RA Sakai K., Okido T., Furuno M., Aono H., Baldarelli R., Barsh G.,
RA Blake J., Boifelli D., Boylston N., Carrinchi P., de Bonaldo M.F.,
RA Brownstein M.J., Bult C., Fletcher C., Fujita M., Gariboldi M.,
RA Gustincich S., Hill D., Hofmann M., Hume D.A., Kamiya M., Lee N.H.,
RA Lyons P., Marchionni L., Mashima J., Mazzarelli J., Mombaerts P.,

```


RESULT 14					
Q9D7C5 PRELIMINARY; PRT: 395 AA.					
ID	C9D7C5;				
AC	Q9D7C5;				
DT	01-JUN-2001 (TrEMBLrel. 17, Created)				
DT	01-JUN-2001 (TrEMBLrel. 17, Last sequence update)				
DE	01-MAR-2003 (TrEMBLrel. 23, Last annotation update)				
DE	2310051B21R1k protein.				
GN	2310051B21R1K.				
OS	Mus musculus (Mouse).				
OC	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;				
OC	Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.				
NCBI_TaxId	=10090;				
RN	[1]				
RP	SEQUENCE FROM N.A.				
RC	MDLINE=C57BL/6J; TISSUE=Tongue;				
RX	STRFAME=21085660; PubMed=11217851;				
RA	Kawai J., Shinagawa A., Shibata K., Yoshino M., Itoh M., Ishii Y.,				
RA	Arakawa T., Hara A., Fukunishi Y., Konno H., Adachi J., Fukuda S.,				
RA	Aizawa K., Izawa M., Nishi K., Kiyoosawa H., Kondo S., Yamanaoka I.,				
RA	Saito T., Okazaki Y., Gotojori T., Bono H., Kasukawa T., Saito R.,				
RA	Kadota K., Matsuda H.A., Ashburner M., Batalov S., Casavant T.,				
RA	Fleischmann W.S., Gaasterland T., Gissi C., King B.G., Kochiya H.,				
RA	Kuehl P., Lewis S., Matsuo Y., Nikaido I., Pesole G., Quackenbush J.,				
RA	Schirml L.M., Staubli F., Suzuki R., Tomita M., Wagner L., Washio T.,				
RA	Sakai K., Okido T., Furuno M., Aono H., Baldarelli R., Barsh G.,				
RA	Blake J., Boffelli D., Bojunga N., Carninci P., de Bonaldo M.F.,				
RA	Brownstein M.J., Bult C., Fletcher C., Fujita M., Gariboldi M.,				
RA	Gustincich S., Hill D., Hofmann M., Hume D.A., Kamiya M., Lee N.H.,				
RA	Lyon P., Marchionni L., Mashima J., Mazzarelli J., Momabato N.,				
RA	Nordone P., Ring B., Ringwald M., Rodriguez I., Sakamoto N.,				
RA	Sasaki H., Sato K., Schoenbach C., Seya T., Shibata Y., Storch K.-F.,				
RA	Suzuki H., Toyooka K., Wang K.H., Weitz C., Whittaker C., Wilming L.,				
RA	Wynshaw-Boris A., Yoshida K., Haegawa Y., Kawaji H., Kohtsuki S.,				
RA	Hayashizaki Y.				
RT	"Functional annotation of a full-length mouse cDNA collection.";				
RL	Nature 409:685-690(2001).				
DR	EMBL; AK009359; BAB26240.1; -.				
DR	HSPG; P07098; IHLG.				
DR	MGD; MG1:1914967; 2310051B21R1k.				
DR	InterPro; IPRO00073; A/b hydrolase.				
DR	InterPro; IPRO00734; Lipase.				
DR	InterPro; IPRO00379; Ser esterase.				
DR	Pfam; PF00561; abhydrolase_i_ site.				
DR	PROSITE; PS00120; LIPASE_SER; 1.				
SQ	SEQUENCE 395 AA; 44748 MW; 3ED43992458DE058 CRC64;				
Query Match 54.1%; Score 1123; DB 11; Length 395;					
Best Local Similarity 55.9%; Pred. No. 3e-95;					
Matches 208; Conservative 63; Mismatches 95; Indels 6; Gaps 2;					
Qy	9 DPFAFNISIIHQGYPCSEYEYATGGVILSNVRIPGLVPQPKTGRPVVLQHGLV 68				
Dd	27 NPEANMVMSQITTWGYPSSEYEVTEDGVILGVIRPYKKNSENIKRPVAYLQHLI 86				
Qy	69 GGASNIWSINLPNNSLGFILADAGPDVVMGNSRGNAWRKHKTLISIDQDFWAFSYDEVAR 128				
Dd	87 ASAKNWITNLPPNNSLAFILADAGVDVWLNGSRGTNRKNVYSPDSVFWEAFSFEVMAK 146				
Qy	129 FDPAPVINFTLOKTQSKIYYVGVSQCTWGTFIAFSMTPELAOKIKVFFALAPTATVKHA 188				
Dd	147 YDLPATIDFIVQTKQBQSKIHVGHSGQTGTGFIASFNPALAKKIKRFIALPVAIVTKYT 206				
Qy	189 KSPGTKEFLLPDMMIKGLFGKKEFL---YQTRFRQLVIYLCGVVIDIQICSNITMLLGG 245				
Dd	207 ESPFKKHFTPKFLKLVIFGNKMFMPHNYLDQFLGTEV--CSRELLDLLCSNALFIPCG 263				
Qy	246 FNTNNMNSRASVYAHTLAGTSYONILHGSQAVNSGELRAFDMGSETNKLEKNQPTFV 305				
Dd	264 FDKKLNLRSDPDVLYGNPNAGTSTQDLFHWAQLAKSGKLQAINMGSPLONMHLHYNQKTPP 323				
Qy	306 RYRVDRMTVPTAMWTGGQDWLSNPEDVKMLSEVNLTLYHKNIPEWAHVDFIWGLDAPH 365				

```
Db 324 YDVSAMTVPVAVNGGHDILADPDQVAMLLPKLPNLLYHKEILPYNHLDFIWMADAPQE 383
QY 366 MYNEIHLMOQE 377
Db 384 VYNEIVTMAED 395

RESULT 15
QSD6L1
ID QSD6L1 PRELIMINARY; PRT; 395 AA.
AC QSD6L1;
DT 01-JUN-2001 (TREMELrel. 17, Created)
DT 01-JUN-2001 (TREMELrel. 17, Last sequence update)
DT 01-MAR-2003 (TREMELrel. 23, Last annotation update)
DE Adult male tongue cDNA, RIKEN full-length enriched library,
DE clone:2310079020, full insert sequence.
GN 2310051B21RIK.
OS Mus musculus (Mouse).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
OC NCBI_TaxID=10090;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=C57BL/6J; TISSUE=Tongue;
RX MEDLINE=21085660; PubMed=11217851;
RA Kawai J., Shinagawa A., Shibata K., Yoshino M., Itoh M., Ishii Y.,
RA Arakawa T., Hara A., Fukunishi Y., Konno H., Adachi J., Fukuda S.,
RA Aizawa K., Izawa M., Nishi K., Kiyosawa H., Kondo S., Yamanaka I.,
RA Saito T., Okazaki Y., Gojobori T., Bono H., Kasukawa T., Saito R.,
RA Kadota K., Matsuda H.A., Ashburner M., Batalov S., Casavant T.,
RA Fleischmann W., Gaasterland T., Glissi C., King B., Kochiwa H.,
RA Kueni P., Lewis S., Matsuo Y., Nikaido I., Pesole G., Quackenbush J.,
RA Schriml L.M., Stabili P., Suzuki R., Tomita M., Wagner L., Washio T.,
RA Sakai K., Okido T., Furuno M., Aono H., Baldarelli R., Barsh G.,
RA Blake J., Boffelli D., Bojunga N., Carninci P., de Bonaudo M.F.,
RA Brownstein M.J., Bult C., Fletcher C., Fujita M., Gariboldi M.,
RA Gustincich S., Hill D., Hofmann M., Hume D.A., Kamiya M., Lee N.H.,
RA Lyons P., Marchionni L., Mashima J., Mazzerelli J., Mombaerts P.,
RA Nordone P., Ring B., Ringwald M., Rodriguez I., Sakamoto N.,
RA Sasaki H., Sato K., Schoenbach C., Seya T., Shibata Y., Storch K.-F.,
RA Suzuki H., Toyooka K., Wang K.H., Weitz C., Whittaker C., Wilming L.,
RA Wynshaw-Boris A., Yoshida K., Hasegawa Y., Kawaji H., Kohsaki S.,
RA Hayashizaki Y.;
RT "Functional annotation of a full-length mouse cDNA collection.";
RL Nature 409:685-690(2001).
DR EMBL; AK010231; BAB26784.1; -.
DR HSSP; P07098; IHG.
DR MGD; MGI:1914967; 2310051B21RIK.
DR InterPro; IPR000073; A/b hydrolase.
DR InterPro; IPR000734; Lipase.
DR InterPro; IPR000379; Ser. esters site.
DR Pfam; PF00561; abhydrolase_1_-site.
DR PROSITE; PS00120; LIPASE_SER; 1.
SQ SEQUENCE 395 AA; 44659 MW; 1350BB95FCSA40E0 CRC64;

Query Match 54.0%; Score 1122; DB 11; Length 395;
Best Local Similarity 56.2%; Pred. No. 3.7e-95;
Matches 209; Conservative 63; Mismatches 94; Indels 6; Gaps 2;

QY 9 DPEAFNISEIIOHGYCEEEVATEGDIYLSVNRIPRGLVQPKYKTSRPPVLLQHLV 69
Db 27 NPEANNVNSQITTYGPFSEEEVVTEDGYILGVTRIPYKKGNSENIGKRFVAYLOHGLI 86
QY 69 GCASWISNLPNSLGLTADAGFDVWNGNRGNASRKHKTLSIDODEFWAFSPDEMAR 128
Db 87 ASATWITNLPNSLAFIADAGYDVWLNSENGNTWSRKNYYSFDSVEFWAFSDEWAK 146
QY 129 FDLPAVINFLQKQOEKIYYGVYSGTTMGFIAPFTMPPELAQKIKMYFALAPIATVKA 188
Db 147 YDLPATIDFIVQKIQEIKHYVGHSGQTTIGFIAPFTNPALAKKIKRFYALAPVATKYT 206
QY 189 KSPGKTKLLPDMMIKGLFGKKEFL--YQTFLEQLVILCGVILDQICSNIWLLGG 245
```

```
Db 207 ESPFKKISLIHKFLKLVIFGNWFWPHNYLDQFLCTEV---CSRELLDLCSNALFICG 263
QY 246 FNTNNNMGRASVYAAHTLAGTSVQNIHLWSQAVNSGELRAPDQWGETYNLEKCMQPTPV 305
Db 264 FDKCNLNVSRFDVYLGHENPAGTSTQDLFEWAQAKSGKLOAYNWGSPLONLHLHYNQKTPP 323
QY 306 RYRVRDMTVPYTAWMTGGQDWLSNPEDVKVLLSEVNTLYHKNIPEWAHVDFIWLGLDAPHR 365
Db 324 YDVSAMTVPVAVNGGHDILADPDQVAMLLPKLPNLLYHKEILPYNHLDFIWMADAPQE 383
QY 366 MYNEIHLMOQE 377
Db 384 VYNEIVTMAED 395

Search completed: February 19, 2004, 14:22:25
Job time : 28.8229 secs
```

GenCore version 5.1.6
Copyright (c) 1993 - 2004 Compugen Ltd.

OM protein - protein search, using sw model

Run on: February 19, 2004, 14:14:34 ; Search time 32.1402 Seconds
(without alignments)
1926.040 Million cell updates/sec

Title: US-10-042-431-49

Perfect score: 2076

Sequence: 1 VHMPTKAVDPAPMNISEII.....IHLMQEETNLQSGRCRAVL 390

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 1107863 seqs, 158726573 residues

Total number of hits satisfying chosen parameters: 1107863

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : A_Geneseq_19Jun03.*
1: /SIDS1/gcgdata/geneseq/geneseq-emb1/AA1980.DAT.*
2: /SIDS1/gcgdata/geneseq/geneseq-emb1/AA1981.DAT.*
3: /SIDS1/gcgdata/geneseq/geneseq-emb1/AA1982.DAT.*
4: /SIDS1/gcgdata/geneseq/geneseq-emb1/AA1983.DAT.*
5: /SIDS1/gcgdata/geneseq/geneseq-emb1/AA1984.DAT.*
6: /SIDS1/gcgdata/geneseq/geneseq-emb1/AA1985.DAT.*
7: /SIDS1/gcgdata/geneseq/geneseq-emb1/AA1986.DAT.*
8: /SIDS1/gcgdata/geneseq/geneseq-emb1/AA1987.DAT.*
9: /SIDS1/gcgdata/geneseq/geneseq-emb1/AA1988.DAT.*
10: /SIDS1/gcgdata/geneseq/geneseq-emb1/AA1989.DAT.*
11: /SIDS1/gcgdata/geneseq/geneseq-emb1/AA1990.DAT.*
12: /SIDS1/gcgdata/geneseq/geneseq-emb1/AA1991.DAT.*
13: /SIDS1/gcgdata/geneseq/geneseq-emb1/AA1992.DAT.*
14: /SIDS1/gcgdata/geneseq/geneseq-emb1/AA1993.DAT.*
15: /SIDS1/gcgdata/geneseq/geneseq-emb1/AA1994.DAT.*
16: /SIDS1/gcgdata/geneseq/geneseq-emb1/AA1995.DAT.*
17: /SIDS1/gcgdata/geneseq/geneseq-emb1/AA1996.DAT.*
18: /SIDS1/gcgdata/geneseq/geneseq-emb1/AA1997.DAT.*
19: /SIDS1/gcgdata/geneseq/geneseq-emb1/AA1998.DAT.*
20: /SIDS1/gcgdata/geneseq/geneseq-emb1/AA1999.DAT.*
21: /SIDS1/gcgdata/geneseq/geneseq-emb1/AA2000.DAT.*
22: /SIDS1/gcgdata/geneseq/geneseq-emb1/AA2001.DAT.*
23: /SIDS1/gcgdata/geneseq/geneseq-emb1/AA2002.DAT.*
24: /SIDS1/gcgdata/geneseq/geneseq-emb1/AA2003.DAT.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	2076	100.0	390	22	Human TANGO 294 ma
2	2076	100.0	409	22	Human CG162 (or C5
3	2076	100.0	423	22	Human TANGO 294.
4	2076	100.0	423	24	Amino acid sequenc
5	2005.5	96.6	423	23	Human lipid-associ
6	1991	95.9	398	23	Human lipase prote
7	1991	95.9	398	23	Human lysosomal ac
8	1536.5	74.0	371	23	Human lysosomal ac
9	1289	62.1	399	22	Human shear stress

10	1289	62.1	399	22	AA866061	Human lysosomal ac
11	1282	61.8	392	23	AA018227	Human lysosomal ac
12	1181.5	56.9	449	22	AAU30498	Novel human secret
13	1165	56.1	221	22	AA866068	Human TANGO 294 ex
14	1153	55.5	395	6	AA850322	Rat lingual lipase
15	1150.5	55.4	398	14	AA837302	RGL precursor. Or
16	1141.5	55.0	379	15	AA856870	Canine gastric lip
17	1141.5	55.0	379	17	AA809382	Dog gastric lipase
18	1141.5	55.0	380	15	AA856871	Canine gastric lip
19	1133	54.6	395	23	AAE14744	Human triacylglyce
20	1133	54.6	399	23	AAU77496	Human lipid metabo
21	1129	54.3	399	22	AA867513	Amino acid sequenc
22	1127	54.3	378	23	AB876189	Kid goat peregastri
23	1126	54.2	397	23	AAE17308	Human lysosomal ac
24	1109	53.4	398	7	AA860724	Sequence of pegas
25	1109	53.4	398	7	AA860658	Sequence of human
26	1109	53.4	398	17	AAW09383	Human gastric lipa
27	1109	53.4	398	22	AA866086	Human lipase prote
28	1109	53.4	398	24	ABP96698	Human gastric lipa
29	1093	52.6	403	23	AAU77493	Human lipid metabo
30	1082	52.1	427	23	AAU98539	Human lysosomal ac
31	1024	49.3	401	23	ABG94641	Human NOV6b protei
32	1023	49.3	395	23	AA018226	Human lysosomal ac
33	1023	49.3	395	24	ABG73144	Human lipase prote
34	998	48.1	373	23	ABP53569	Human NOV1 protei
35	995.5	48.0	362	24	ABP96697	Human triacylglyce
36	969	46.7	365	23	AAE17307	Human lysosomal ac
37	909	43.8	331	24	ABR41626	Human DITHP secret
38	893	43.0	731	23	AAE14746	Human triacylglyce
39	890.5	42.9	390	23	ABG94640	Human NOV6a protei
40	783	37.7	144	22	AA866070	Human TANGO 294 cy
41	767	36.9	276	23	AAE14745	Human triacylglyce
42	767	36.9	289	23	AAU77494	Human lipid metabo
43	731.5	35.2	280	23	AAE14743	Human triacylglyce
44	710.5	34.2	656	22	ABG26839	Novel human diagno
45	681	32.8	233	22	AA861608	Human protein HP03

ALIGNMENTS

RESULT 1

ID AA866067 standard; Protein; 390 AA.

AC AA866067;

DT 30-MAR-2001 (first entry)

DE Human TANGO 294 mature protein.

KW TANGO protein; INTERCEPT protein; neurological disorder;

KW central nervous system; focal brain disorder; bipolar affective disorder;

KW global-diffuse cerebral disorder; cerebrovascular; Alzheimer's disease;

KW senile dementia; Huntington's disease; amyotrophic lateral sclerosis;

KW Parkinson's; Gilles de la Tourette's syndrome; hypertension; sleep;

KW neuropsychiatric; psychoactive substance use; anxiety.

OS Homo sapiens.

PN WO2000077239-A2.

XX 21-DEC-2000.

XX 24-MAY-2000; 2000WO-US14858.

XX 14-JUN-1999; 99US-0333159.

XX (MILL-) MILLENNIUM PHARM INC.

PI McCarthy SA, Fraser CC, Sharp JD, Barnes TM;

WPI; 2001-032313/04.

XX TANGO and INTERCEPT nucleic acids, proteins, and antibodies, useful for
PT screening assays and diagnostic assays and for the treatment of
PT neurological diseases such as Alzheimer's, Parkinson's and Huntington's
PT disease -
XX
XX
PS Claim 8; Pages 324-325; 359pp; English.
XX
XX The present invention relates to TANGO or INTERCEPT proteins and coding
CC sequences (see AAF45121-F45136 and AAF45138-F45139 and AAB66031-B66057,
CC AAB66064-B66083 and AAB66085). The TANGO/INTERCEPT proteins and coding
CC sequences are useful for the treatment of neurological disorders such as
CC central nervous system (CNS) disorders, CNS-related disorders, focal
CC brain disorders, global-diffuse cerebral disorders and other
CC neurological and cerebrovascular disorders. The CNS disorders include
CC Alzheimer's disease, senile dementia, Huntington's disease, amyotrophic
CC lateral sclerosis, Parkinson's, Gilles de la Tourette's syndrome,
CC autonomic function disorders such as hypertension and sleep disorders,
CC neuropsychiatric disorders, psychoactive substance use disorders,
CC anxiety, and bipolar affective disorder.
XX
XX Sequence 390 AA;

Query Match 100.0%; Score 2076; DB 22; Length 390;
Best Local Similarity 100.0%; Pred. No. 1.4e-204;
Matches 390; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 1 VHMPTKAVDPEAFMNISEIIHQGYPCBEYEVATEDGYILSVNIRPGLVQPKTGRPV 60
Db 1 VHMPTKAVDPEAFMNISEIIHQGYPCBEYEVATEDGYILSVNIRPGLVQPKTGRPV 60
QY 61 VLLQHLGVGASNWIINLPNNISLGFILADAGFDVWVGNRGNWSRGNWSRKHKTLSIDQDEFWA 120
Db 61 VLLQHLGVGASNWIINLPNNISLGFILADAGFDVWVGNRGNWSRGNWSRKHKTLSIDQDEFWA 120
QY 121 PSYDEMARFDPVAVINFLQKTQGEKIYVGVYSGTWTGFTAFSTMPFLAQKIMYFALA 180
Db 121 PSYDEMARFDPVAVINFLQKTQGEKIYVGVYSGTWTGFTAFSTMPFLAQKIMYFALA 180
QY 181 PIATVHAKSPGTFKLLPDMNKGFLGKKEFLYQTRFLQVLYLQCGVILDCQINIM 240
Db 181 PIATVHAKSPGTFKLLPDMNKGFLGKKEFLYQTRFLQVLYLQCGVILDCQINIM 240
QY 241 LLLGGNTNNMNSRVSVAHTLACTSVQNTLHWSQVNSGELPAFQWSETKNLEKN 300
Db 241 LLLGGNTNNMNSRVSVAHTLACTSVQNTLHWSQVNSGELPAFQWSETKNLEKN 300
QY 301 QPTPVRYRVRDVTPTAMWTGGQDMLSNPDEYKMLLSEVTNLIYHKNIPENAHVDPIWGL 360
Db 301 QPTPVRYRVRDVTPTAMWTGGQDMLSNPDEYKMLLSEVTNLIYHKNIPENAHVDPIWGL 360
QY 361 DAPHRMNEIHLMOQEEITLSQGRCEAVL 390
Db 361 DAPHRMNEIHLMOQEEITLSQGRCEAVL 390

RESULT 2
AAE11931
ID AAE11931 standard; Protein; 409 AA.
XX
XX AAE11931;
XX
XX 18-DEC-2001 (first entry)
XX
XX Human CG162 (or C59) lipase protein #2.
XX
XX Human; apolipoprotein; lipase; lipoprotein receptor; ALLr; angina;
KW cardiovascular disease; lipid metabolism; myocardial infarction;
KW cerebral ischaemia; arterial thrombosis; thrombolytic; antilipemic;
KW coronary artery thrombosis; cerebral artery thrombosis; stroke;
KW intracardiac thrombosis; gene therapy; cardiovascular; vasodilator;
KW neuroprotectant; cerebroprotective.
XX

OS Homo sapiens.
XX Key Location/Qualifiers
FH Peptide 1..19
FT Protein /label= Signal_peptide
FT 20..409
FT /note= "Human mature CG162 (or C59) lipase protein"
XX
XX WO200179446-A2.
XX 25-OCT-2001.
XX 16-APR-2001; 2001WO-US12529.
XX 14-APR-2000; 2000US-197137P.
XX 20-JUN-2000; 2000US-0598042.
XX 03-AUG-2000; 2000US-0631451.
XX 22-SEP-2000; 2000US-0667298.
XX 17-NOV-2000; 2000US-0714936.
XX (HYSB-) HYSBQ INC.
XX
XX Ballinger DG, Loeb D, Montgomery JR, Tang TY, Zhou P, Goodrich R;
PI Liu C, Asundi V, Zhao QA, Wehrman T, Drmanac RT, Ren F, Qian XB;
PI Wang D;
XX
XX WPI; 2001-611724/70.
XX N-PSDB; AAD19226.
XX
XX Nucleic acids encoding human apolipoproteins, lipases, and lipoprotein
PT receptor polypeptides, useful for preventing diagnosing and treating
PT lipid metabolism disorders, thrombosis and cardiovascular diseases -
XX
XX Claim 10; Fig 3; 266pp; English.

The invention relates to polynucleotides encoding proteins CG122, CG179,
CG95, CG121, CG162, CG27, CG153 and CG168 which are related to proteins
involved in lipid metabolism and cardiovascular disease such as human
apolipoproteins, lipases and lipoprotein receptor proteins. These DNA
and protein sequences are useful for treating or preventing disorders
associated with apolipoproteins, lipases and lipoprotein receptor (ALLr)
expression and for treating lipid metabolism, cardiovascular diseases
and thrombosis. Antibodies against these proteins are useful for
determining the presence of or predisposition to a disease associated
with altered levels of these sequences. ALLr polypeptides are also
useful for identifying agents (agonists and antagonists) that bind to
them and cells expressing ALLr proteins are useful for identifying a
therapeutic agent for use in treatment of a pathology related to
aberrant expression or physiological interactions of this polypeptide.
Vectors comprising these DNA and protein sequences are also useful for
producing ALLr proteins. The nucleic acids and polypeptides of the
invention are also useful for the treatment of occlusive cardiovascular
diseases, myocardial infarction, cerebral ischaemia, angina, arterial
thrombosis, coronary artery thrombosis and cerebral artery thrombosis
or intracardiac thrombosis and stroke. The nucleotides of the invention
are used in gene therapy. The present sequence is human CG162 (or C59)
lipase protein.

Sequence 409 AA;
Query Match 100.0%; Score 2076; DB 22; Length 409;
Best Local Similarity 100.0%; Pred. No. 1.5e-204;
Matches 390; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 1 VHMPTKAVDPEAFMNISEIIHQGYPCBEYEVATEDGYILSVNIRPGLVQPKTGRPV 60
Db 20 VHMPTKAVDPEAFMNISEIIHQGYPCBEYEVATEDGYILSVNIRPGLVQPKTGRPV 79
QY 61 VLLQHLGVGASNWIINLPNNISLGFILADAGFDVWVGNRGNWSRGNWSRKHKTLSIDQDEFWA 120
Db 80 VLLQHLGVGASNWIINLPNNISLGFILADAGFDVWVGNRGNWSRGNWSRKHKTLSIDQDEFWA 139
QY 121 PSYDEMARFDPVAVINFLQKTQGEKIYVGVYSGTWTGFTAFSTMPFLAQKIMYFALA 180

Db 140 PSYDEWARFDPVAVINFLQKTQGEKIYYVGSQGTMTGFIATFSTMPPELAQKIKMYFALA 199
 QY 181 PIATVKHAKSPGTFKLLPDMKIKGLFGKKEFLYQTRFLRQLVLYLGGVILDOICSNIM 240
 Db 200 PIATVKHAKSPGTFKLLPDMKIKGLFGKKEFLYQTRFLRQLVLYLGGVILDOICSNIM 259
 QY 241 LLLGGFNTNNMMSRASVYAAHTLAGTSVQNILHWSQAVNSGELRAPDWGSETKNLEKCN 300
 Db 260 LLLGGFNTNNMMSRASVYAAHTLAGTSVQNILHWSQAVNSGELRAPDWGSETKNLEKCN 319
 QY 301 QPTPVRYRVDMTVPTAMWTGGQDWLSNPEDVKMLLSEVTNLIYHKNIPEWAHVDFPIWGL 360
 Db 320 QPTPVRYRVDMTVPTAMWTGGQDWLSNPEDVKMLLSEVTNLIYHKNIPEWAHVDFPIWGL 379
 QY 361 DAPHRMYNEIHLHMQOEETNLSQGRCEAVL 390
 Db 380 DAPHRMYNEIHLHMQOEETNLSQGRCEAVL 409

RESULT 3

AA08369
 ID AAB66065 standard; Protein; 423 AA.

XX
 AC AAB66065;

XX
 DT 30-MAR-2001 (first entry)

XX
 DE Human TANGO 294.

XX
 KW TANGO protein; INTERCEPT protein; neurological disorder;

XX
 KW central nervous system; focal brain disorder; bipolar affective disorder;
 KW global diffuse cerebral disorder; cerebrovascular; Alzheimer's disease;
 KW senile dementia; Huntington's disease; amyotrophic lateral sclerosis;
 KW Parkinson's; Gilles de la Tourette's syndrome; hypertension; sleep;
 KW neuropsychiatric; psychoactive substance use; anxiety.

XX
 OS Homo sapiens.

XX
 PN WO200077239-A2.

XX
 PD 21-DEC-2000.

XX
 PF 24-MAY-2000; 2000WO-US14858.

XX
 PR 14-JUN-1999; 99US-0333159.

XX
 PA (MILL-) MILLENNIUM PHARM INC.

XX
 PI McCarthy SA, Fraser CC, Sharp JD, Barnes TM;

XX
 DR WPI; 2001-032313/04.

XX
 DR N-PSDB; AAP45131, AAP45132.

XX
 PT TANGO and INTERCEPT nucleic acids, proteins, and antibodies, useful for
 PT screening assays and diagnostic assays and for the treatment of
 PT neurological diseases such as Alzheimer's, Parkinson's and Huntington's
 PT disease -

XX
 PS Claim 8; Fig 6; 359pp; English.

XX
 CC The present invention relates to TANGO or INTERCEPT proteins and coding
 CC sequences (see AAP45121-F45136 and AAP45138-F45139 and AAB66031-B66057,
 CC AAB66064-B66083 and AAB66085). The TANGO/INTERCEPT proteins and coding
 CC sequences are useful for the treatment of neurological disorders such as
 CC central nervous system (CNS) disorders, CNS-related disorders, focal
 CC brain disorders, global-diffuse cerebral disorders and other
 CC neurological and cerebrovascular disorders. The CNS disorders include
 CC Alzheimer's disease, senile dementia, Huntington's disease, amyotrophic
 CC lateral sclerosis, Parkinson's, Gilles de la Tourette's syndrome,
 CC autonomic function disorders such as hypertension and sleep disorders,
 CC neuropsychiatric disorders, psychoactive substance use disorders,
 CC anxiety, and bipolar affective disorder.

XX
 SQ Sequence 423 AA;

Query Match 100.0%; Score 2076; DB 22; Length 423;
 Best Local Similarity 100.0%; Pred. No. 1.6e-204;
 Matches 390; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 VHMPTKAVDPEAFPMNISEIIHQGYPCBEYEVATEDGYILSVNRIPRGLVQPKKTGSRPV 60
 Db 34 VHMPTKAVDPEAFPMNISEIIHQGYPCBEYEVATEDGYILSVNRIPRGLVQPKKTGSRPV 93
 QY 61 VLLQHLVGGASNWSINLNNLSGIFILADAGPDVVMGNSRGNANRKHKTLSIDQDEFWA 120
 Db 94 VLLQHLVGGASNWSINLNNLSGIFILADAGPDVVMGNSRGNANRKHKTLSIDQDEFWA 153
 QY 121 FSYDEWARFDPVAVINFLQKTQGEKIYYVGSQGTMTGFIATFSTMPPELAQKIKMYFALA 180
 Db 154 FSYDEWARFDPVAVINFLQKTQGEKIYYVGSQGTMTGFIATFSTMPPELAQKIKMYFALA 213
 QY 181 PIATVKHAKSPGTFKLLPDMKIKGLFGKKEFLYQTRFLRQLVLYLGGVILDOICSNIM 240
 Db 214 PIATVKHAKSPGTFKLLPDMKIKGLFGKKEFLYQTRFLRQLVLYLGGVILDOICSNIM 273
 QY 241 LLLGGFNTNNMMSRASVYAAHTLAGTSVQNILHWSQAVNSGELRAPDWGSETKNLEKCN 300
 Db 274 LLLGGFNTNNMMSRASVYAAHTLAGTSVQNILHWSQAVNSGELRAPDWGSETKNLEKCN 333
 QY 301 QPTPVRYRVDMTVPTAMWTGGQDWLSNPEDVKMLLSEVTNLIYHKNIPEWAHVDFPIWGL 360
 Db 334 QPTPVRYRVDMTVPTAMWTGGQDWLSNPEDVKMLLSEVTNLIYHKNIPEWAHVDFPIWGL 393
 QY 361 DAPHRMYNEIHLHMQOEETNLSQGRCEAVL 390
 Db 394 DAPHRMYNEIHLHMQOEETNLSQGRCEAVL 423

RESULT 4

ABU08369
 ID ABU08369 standard; Protein; 423 AA.

XX
 AC ABU08369;

XX
 DT 03-JUN-2003 (first entry)

XX
 DE Amino acid sequence for human TANGO 294.

XX
 KW Human; TANGO 294; INTERCEPT; cellular process; tissue typing;
 KW forensic biology; cellular protease activity; cell interaction;
 KW development; blood disorder; haematopoietic cell-related disorder;
 KW growth; cell proliferation; cell differentiation; gamma delta T-cell;
 KW immune system; metabolic disorder; homeostatic disorder; anaemia;
 KW developmental bone disorder; osteoporosis; bacterial infection;
 KW thrombocytopaenia; renal failure; liver disease; cystic fibrosis;
 KW obesity; cancer; bronchitis; asthma; emphysema; pulmonary oedema;
 KW respiratory distress syndrome; osteopathic; antibacterial; antianemic;
 KW thrombolytic; nephrotropic; antiobesity; hepatotropic; cytostatic;
 KW antiinflammatory; antiasthmatic.

XX
 OS Homo sapiens.

XX
 FH Key Location/Qualifiers
 FT 1.33

FT /label= Predicted signal peptide
 FT /note= "Given as SEQ ID No:48 and specifically
 FT Claimed in Claim 8"

FT Protein 34..423

FT /label= Predicted mature TANGO_294 protein
 FT /note= "Given as SEQ ID No:49 and specifically
 FT Claimed in Claim 8"

FT Domain 34..254

FT /label= Extracellular domain
 FT /note= "Given as SEQ ID No:50 and specifically
 FT Claimed in Claim 8"

(I). (I), (II) and antibody to (II) are useful for treating a disease or condition associated with decreased or increased expression of functional LIPAM. (I) or (II) is useful for diagnosing, treating or preventing cancer e.g. leukemia, cancer of adrenal gland, bladder, bone, bone marrow, brain, ovary, etc; cardiovascular disorders e.g. atherosclerosis, hypertension, Raynaud's disease, angina pectoris, myocardial infarction, rheumatic fever, cardiomyopathy, pericarditis, chronic obstructive pulmonary disease, etc; neurological disorders e.g. Alzheimer's disease, stroke, epilepsy, dementia, Parkinson's disease, multiple sclerosis, Creutzfeldt-Jakob disease, myasthenia gravis, anxiety, diabetic neuropathy, etc; autoimmune/inflammatory disorders e.g. acquired immunodeficiency syndrome (AIDS), Addison's disease, anemia, asthma, bronchitis, diabetes mellitus, emphysema, Crohn's disease, atopic dermatitis, rheumatoid arthritis, psoriasis, systemic lupus erythematosus, trauma, etc; and gastrointestinal disorders e.g. gastritis, anorexia, nausea gastroenteritis, ulcerative colitis, cholecystitis, hepatitis, cirrhosis, hepatoma, diarrhoea, constipation, etc, and disorders of lipid metabolism e.g., fatty liver, cholestasis, Fabry's disease, hypercholesterolaemia. (I) or (II) is useful for assessing the effects of exogenous compounds on the expression of (I) or (II). (I) is useful in tissue or cell type. (I) or antibody to (I) is useful as elements on a microarray. (II) is useful for creating knockin humanized animals or transgenic animals to model human diseases, in somatic or germline gene therapy, to generate a transcript image of a tissue or cell type, for detecting differences in the chromosomal location due to translocation, inversion, etc., among normal, carrier or affected individuals, and as hybridisation probes for mapping naturally occurring genomic sequences. The present sequence represents the amino acid sequence of a human lipid-associated molecule (LIPAM) of the invention.

Sequence 423 AA;

Query Match 96.6%; Score 2005.5; DB 23; Length 423;
Best Local Similarity 97.2%; Pred. No. 2.8e-197;
Matches 379; Conservative 2; Mismatches 2; Indels 7; Gaps 1;
QY 8 VDPAFNISEIIHQGPCEEEVATEDGVLSVNIIPGLVQPKTKGSRPVLLQHL 67
Db 34 VNPEVWNTSEIIHQGPCEEEVATEDGVLSVNIIPGLVQPKTKGSRPVLLQHL 93
QY 68 VGGASNIISLNNLSGLFADAGFDVWNGSRGNAMSRKHTLSIDQDFWAFSYDEVA 127
Db 94 VGGASNIISLNNLSGLFADAGFDVWNGSRGNAMSRKHTLSIDQDFWAFSYDEVA 153
QY 128 RFDLPVAVINFLQTKGQEKIYYGYSGQTTMGPTAFSTMPELAQIKMYPALAPIATVKH 197
Db 154 RFDLPVAVINFLQTKGQEKIYYGYSGQTTMGPTAFSTMPELAQIKMYPALAPIATVKH 213
QY 188 AKSPGTFKLLLPDMWIKGLFKKEFLYQTRFLQVLYLQGVLLDQICSNIMLLGGFN 247
Db 214 AKSPGTFKLLLPDMWIKGLFKKEFLYQTRFLQVLYLQGVLLDQICSNIMLLGGFN 273
QY 248 TNNMNM-----SRASYAHTLAGTSVQNILHWSQAVNSGELRAPDWGSETNLEKCN 300
Db 274 TNNMNMTHGLQSRASYAHTLAGTSVQNILHWSQAVNSGELRAPDWGSETNLEKCN 333
QY 301 QPTPVYRVVDMVTPTAMTGGQDLNPNEDVKMLSEVTNLIYHKNIPEWAHVDFIWL 360
Db 334 QPTPVYRVVDMVTPTAMTGGQDLNPNEDVKMLSEVTNLIYHKNIPEWAHVDFIWL 393
QY 361 DAPHRMNEIHLMQBETNLSQGRCEAVL 390
Db 394 DAPHRMNEIHLMQBETNLSQGRCEAVL 423

RESULT 6
ABG71000
ID ABG71000 standard; Protein; 398 AA.
XX AC
XX ABG71000;

DT 07-JAN-2003 (first entry)
XX Human lipase protein.
DE Human; lipase; cancer; colon; kidney; skin; brain; testis; chromosome 10;
KW Burkitt's lymphoma; cytostatic; gene therapy; lipase-inhibitor.
XX Homo sapiens.
XX Key Location/Qualifiers
FT Modified-site 88..91 /note= "N-glycosylation site"
FT Modified-site 260..263 /note= "N-glycosylation site"
FT Modified-site 388..391 /note= "N-glycosylation site"
FT Modified-site 22..25 /note= "cAMP- and cGMP-dependent protein kinase phosphorylation site"
FT Modified-site 113..115 /note= "Protein kinase C phosphorylation site"
FT Modified-site 192..194 /note= "Protein kinase C phosphorylation site"
FT Modified-site 130..133 /note= "Casein kinase II phosphorylation site"
FT Modified-site 150..153 /note= "Casein kinase II phosphorylation site"
FT Modified-site 173..176 /note= "Casein kinase II phosphorylation site"
FT Modified-site 335..338 /note= "Casein kinase II phosphorylation site"
FT Modified-site 149..157 /note= "Casein kinase II phosphorylation site"
FT Modified-site 74..79 /note= "Tyrosine kinase phosphorylation site"
FT Modified-site 77..82 /note= "N-myristoylation site"
FT Modified-site 105..110 /note= "N-myristoylation site"
FT Modified-site 109..114 /note= "N-myristoylation site"
FT Modified-site 163..168 /note= "N-myristoylation site"
FT Modified-site 252..257 /note= "N-myristoylation site"
FT Modified-site 215..218 /note= "Amidation site"
FT Active-site 155..164 /note= "Lipases, serine active site"
XX WO2002/4975-A2.
XX 26-SEP-2002.
XX 18-MAR-2002; 2002WO-US08035.
XX 20-MAR-2001; 2001US-0811825.
XX (PEKE) PE CORP NY.
XX Yan C, Di Francesco V, Beasley EM;
XX WPI; 2002-750560/81.
XX N-PSDB; ABS55356, ABS55357.
XX New isolated human lipase peptides and encoding nucleic acids, useful
XX for diagnosing and treating disorders mediated by human lipase
XX proteins, e.g. cancer of the colon, kidneys, skin and testis, and
XX Burkitt's lymphoma -
XX Claim 1; Fig 2; 79pp; English.
XX The present invention relates to a new lipase peptide. The methods and

compositions of the present invention are useful for diagnosing and treating disorders mediated by the human lipase protein, such as cancer of the colon, kidneys, skin, brain and testis, and Burkitt's lymphoma. The present amino acid sequence represents the human lipase protein of the invention. This sequence is encoded by the human lipase gene located on chromosome 10.

XX
SQ Sequence 398 AA;

Query Match 95.9%; Score 1991; DB 23; Length 398;
Best Local Similarity 100.0%; Pred. No. 8e-196;
Matches 374; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 17 SEIIHQGYPCBEYEYVATEDGYILSVNRIIPRGLVQPKTGSRPVLLQHLVGGASNNIS 76
DB 25 SEIIHQGYPCBEYEYVATEDGYILSVNRIIPRGLVQPKTGSRPVLLQHLVGGASNNIS 84

QY 77 NLPNNSLGFLADAGFDVWNGSRGNWAKSKKHTLSIDQDEFWAFSYDEMAFDFLPVAVIN 136
DB 85 NLPNNSLGFLADAGFDVWNGSRGNWAKSKKHTLSIDQDEFWAFSYDEMAFDFLPVAVIN 144

QY 137 FILQKTQGEKIYYVGYSGQTTMGFIAPSTMPPELAQKIKMYFALAPIATVGHAKSPGPKFL 196
DB 145 FILQKTQGEKIYYVGYSGQTTMGFIAPSTMPPELAQKIKMYFALAPIATVGHAKSPGPKFL 204

QY 197 LLPDMMIKGLFGKKEFLYQTRFLRQLVYLICGVILDDQICSNIMLLGGFNTNNMNSRA 256
DB 205 LLPDMMIKGLFGKKEFLYQTRFLRQLVYLICGVILDDQICSNIMLLGGFNTNNMNSRA 264

QY 257 SVYAAHTLAGTSVQNTILHWSQAVNSGELRAFDFWGSSTKLEKCNQPTPVYRVVDMTVPT 316
DB 265 SVYAAHTLAGTSVQNTILHWSQAVNSGELRAFDFWGSSTKLEKCNQPTPVYRVVDMTVPT 324

QY 317 AMWTGGQDMLSNPDVYKMLLSEVTNLIYHKNIPFWAHVDFIWLGDAPHMYNEIHLMOQ 376
DB 325 AMWTGGQDMLSNPDVYKMLLSEVTNLIYHKNIPFWAHVDFIWLGDAPHMYNEIHLMOQ 384

QY 377 EETNLSQRCCEAVL 390
DB 385 EETNLSQRCCEAVL 398

RESULT 7
AAU99164
ID AAU99164 standard; Protein; 398 AA.
AC AAU99164;
DE 24-SEP-2002 (first entry)
XX Human lysosomal acid lipase #1.
XX Human; enzyme; lysosomal acid lipase; lipid malabsorption illness;
XX cystic fibrosis; alcoholism; heart disease; heart attack;
XX Wolman disease; cholesterol ester storage disease; brain injury;
XX mood disorder; anxiety disorder; thought disorder; volition disorder;
XX sleep disorder; neurogenic disorder; myopathic disorder; COPD;
XX obesity; cancer; chronic obstructive pulmonary disease; diabetes;
XX cardiovascular disorder; Alzheimer's disease; Parkinson's disease;
XX anorexia; osteoarthritis; central nervous system disorder;
XX peripheral nervous system disorder.
XX Homo sapiens.
XX WO200236731-A2.
XX 10-MAY-2002.
XX 30-OCT-2001; 2001WO-EP12518.
XX 31-OCT-2000; 2000US-244170P.
XX 29-MAY-2001; 2001US-293516P.
XX

(PARB) BAYER AG.
XX
FI Xiao Y;
XX WPI; 2002-519248/55.
DR N-PSDB; ABK86569.
XX
XX Novel human lysosomal acid lipase polypeptide, useful for treating cancer, diabetes, obesity, chronic obstructive pulmonary disease, peripheral or central nervous system disorder or cardiovascular disorder

Claim 25; Fig 2; 126pp; English.

The invention relates to a purified human lysosomal acid lipase polypeptide. Also included are the polynucleotide encoding the lipase (or its fragment, derivative, allele or sequence at least 60% identical to it), vectors, host cells, a reagent (e.g. an antisense oligonucleotide) which binds to the lipase or polynucleotide (used for detection and modulating/reducing the lipase activity) and an anti-lipase antibody. The lipase and polynucleotide are useful for identifying therapeutic agents that either increase or decrease the lipase activity. The identified agent, the lipase and polynucleotide are useful for treatment of a disease such as lipid malabsorption illness, cystic fibrosis, alcoholism, heart disease, heart attack, Wolman disease, cholesterol ester storage disease, brain injury, mood disorder, anxiety disorder, thought disorder, volition disorder, sleep disorder, neurogenic disorder, myopathic disorder, obesity, cancer, chronic obstructive pulmonary disease (COPD), diabetes, cardiovascular disorder, Alzheimer's disease, Parkinson's disease, anorexia, osteoarthritis, a central nervous system disorder and a peripheral nervous system disorder. The present sequence is the human lysosomal lipase #1.

SQ Sequence 398 AA;

Query Match 95.9%; Score 1991; DB 23; Length 398;
Best Local Similarity 100.0%; Pred. No. 8e-196;
Matches 374; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 17 SEIIHQGYPCBEYEYVATEDGYILSVNRIIPRGLVQPKTGSRPVLLQHLVGGASNNIS 76
DB 25 SEIIHQGYPCBEYEYVATEDGYILSVNRIIPRGLVQPKTGSRPVLLQHLVGGASNNIS 84

QY 77 NLPNNSLGFLADAGFDVWNGSRGNWAKSKKHTLSIDQDEFWAFSYDEMAFDFLPVAVIN 136
DB 85 NLPNNSLGFLADAGFDVWNGSRGNWAKSKKHTLSIDQDEFWAFSYDEMAFDFLPVAVIN 144

QY 137 FILQKTQGEKIYYVGYSGQTTMGFIAPSTMPPELAQKIKMYFALAPIATVGHAKSPGPKFL 196
DB 145 FILQKTQGEKIYYVGYSGQTTMGFIAPSTMPPELAQKIKMYFALAPIATVGHAKSPGPKFL 204

QY 197 LLPDMMIKGLFGKKEFLYQTRFLRQLVYLICGVILDDQICSNIMLLGGFNTNNMNSRA 256
DB 205 LLPDMMIKGLFGKKEFLYQTRFLRQLVYLICGVILDDQICSNIMLLGGFNTNNMNSRA 264

QY 257 SVYAAHTLAGTSVQNTILHWSQAVNSGELRAFDFWGSSTKLEKCNQPTPVYRVVDMTVPT 316
DB 265 SVYAAHTLAGTSVQNTILHWSQAVNSGELRAFDFWGSSTKLEKCNQPTPVYRVVDMTVPT 324

QY 317 AMWTGGQDMLSNPDVYKMLLSEVTNLIYHKNIPFWAHVDFIWLGDAPHMYNEIHLMOQ 376
DB 325 AMWTGGQDMLSNPDVYKMLLSEVTNLIYHKNIPFWAHVDFIWLGDAPHMYNEIHLMOQ 384

QY 377 EETNLSQRCCEAVL 390
DB 385 EETNLSQRCCEAVL 398

RESULT 8
AAU99165
ID AAU99165 standard; Protein; 371 AA.
XX

AAU99165;
24-SEP-2002 (first entry)
Human lysosomal acid lipase #2.
Human; enzyme; lysosomal acid lipase; lipid malabsorption illness;
cystic fibrosis; alcoholism; heart disease; heart attack;
Wolman disease; cholesterol ester storage disease; brain injury;
mood disorder; anxiety disorder; thought disorder; volition disorder;
sleep disorder; neurogenic disorder; myopathic disorder; COPD;
obesity; cancer; chronic obstructive pulmonary disease; diabetes;
cardiovascular disorder; Alzheimer's disease; Parkinson's disease;
anorexia; osteoarthritis; central nervous system disorder;
peripheral nervous system disorder.
Homo sapiens.
WO200236731-A2.
10-MAY-2002.
30-OCT-2001; 2001WO-EPI2518.
31-OCT-2000; 2000US-244170P.
29-MAY-2001; 2001US-293516P.
(FARB) BAYER AG.
Xiao Y;
WPI; 2002-519248/55.
N-PSDB; ASK6570.
Novel human lysosomal acid lipase polypeptide, useful for treating
cancer, diabetes, obesity, chronic obstructive pulmonary disease,
peripheral or central nervous system disorder or cardiovascular
disorder.
Claim 25; Fig. 5; 126pp; English.
The invention relates to a purified human lysosomal acid lipase
polypeptide. Also included are the polynucleotide encoding the
lipase (or its fragment, derivative, allele or sequence at least 60%
identical to it), vectors, host cells, a reagent (e.g. an antisense
oligonucleotide) which binds to the lipase or polynucleotide (used for
detection and modulating/reducing the lipase activity) and an anti-lipase
antibody. The lipase and polynucleotide are useful for identifying
therapeutic agents that either increase or decrease the lipase activity.
The identified agent, the lipase and polynucleotide are useful for
treatment of a disease such as lipid malabsorption illness,
cystic fibrosis, alcoholism, heart disease, heart attack,
Wolman disease, cholesterol ester storage disease, brain injury,
mood disorder, anxiety disorder, thought disorder, volition disorder,
sleep disorder, neurogenic disorder, myopathic disorder, obesity, cancer,
chronic obstructive pulmonary disease (COPD), diabetes,
cardiovascular disorder, Alzheimer's disease, Parkinson's disease,
anorexia, osteoarthritis, a central nervous system disorder and
a peripheral nervous system disorder. The present sequence is the
human lysosomal lipase #2.
Sequence 371 AA;
Query Match 74.0%; Score 1536.5; DB 23; Length 371;
Best Local Similarity 80.9%; Pred. NO. 4.2e-149;
Matches 297; Conservative 17; Mismatches 22; Indels 31; Gaps 2;
17 SEITQHOGYPCBEVEVATEDGYILSVNRIPLGLVOPKKTGSRPVVILHGLVGCASNWIS 76
25 SEITQHOGYPCBEVEVATEDGYILSVNRIPLGLVOPKKTGSRPVVILHGLVGCASNWIS 84
77 NLPNNSLGFLILADAGFDVWNGSRGNAWSRKHKTLSIDQDEFWAFSYDENARFDPALPAVIN 136

85 NLPNNSLGFLILADAGFDVWNGSRGNAWSRKHKTLSIDQDEFWAFSYDENARFDPALPAVIN 144
137 FILQKTGOEKIYYGYSGQTTMGFIAPSTWPELAQKIMYFALAPIATVVKHAKSPGKFL 196
145 FILQKTGOEKIYYGYSGQTTMGFIAPSTWPELAQKIMYFALAPIATVVKHAKSPGKFL 204
197 LLPDMWIKGLFGKKEFLYQTRFLQRLVYLCGVILQICSNIMLLGGFNNTNNMMN 253
205 LLPDMWIKGLFGKKEFLYQTRFLQRLVYLCGVILQICSNIMLLGGFNNTNNMMN 264
254 -----SRASYAAHTLAGTSVQNIILHWSQAVNSGELRAFDWGSFTKNLEKCNQPTPVYRV 309
265 GLOSRSASYAAHTLAGTSVQNIILHWSQAVNSGELRAFDWGSFTKNLEKCNQPTPVYRV 300
310 RDMVTPTAMWTGGQDMLSNPDEVMALLSEVNTLIYHKNIPDEWHDVFIWGLDAPHRMYNE 369
301 TKIEVPTAIWNGGQDIVADPKDVENLLPQIANLIYKLIPIHYNHVDVFLGSDAQEIYQD 360
370 IHLMOQ 376
361 LILMEE 367
RESULT 9
AAB90783
ID AAB90783 standard; Protein; 399 AA.
XX
AC AAB90783;
XX
DT 15-JUN-2001 (first entry)
XX
DE Human shear stress-response protein SEQ ID NO: 66.
XX
KW Human; shear stress-response protein; vascular disease;
arteriosclerosis.
XX
OS Homo sapiens.
XX
PN WO200125427-A1.
XX
PD 12-APR-2001.
XX
PF 02-OCT-2000; 2000WO-JP06840.
XX
PR 01-OCT-1999; 99JP-0280976.
XX
PA (KIOW) KYOWA HAKKO KOGYO KK.
XX
PA (NOUJ) NOUJIMA H.
XX
PI Nojima H, Yoshigae H, Obayashi M, Ota T, Kawabata A, Sakurada K;
Kuga T, Sekine S, Nakamura Y, Sugano S;
XX
WPI; 2001-266308/27.
DR N-PSDB; AAH02906.
XX
PT DNA sequences, proteins encoded by them and antibodies against them
useful in diagnosis and treatment of vascular disease caused by
arteriosclerosis.
XX
PS Claim 60; Page 402-404; 678pp; Japanese.
XX
CC The present invention provides the protein and coding sequences of a
number of human shear stress response proteins. These are useful in the
CC diagnosis, treatment and screening of vascular diseases caused by
CC arteriosclerosis, including heart failure, post-PTCA restenosis and
CC hypertension.
XX
SQ Sequence 399 AA;
Query Match 62.1%; Score 1289; DB 22; Length 399;
Best Local Similarity 63.5%; Pred. NO. 1.4e-123;
Matches 235; Conservative 55; Mismatches 80; Indels 0; Gaps 0;

QY 7 AVDPFAFMTSEIIHQGYPCPEEYEVATEDGYILSVNRIPLGVLPKKTGSRPVVLLQHG 66
 DB 28 AVDPETNMVSEIIISYMGFPSEYLVETEDGYILCLNRIPLHGRKXSDKGPFPVFLQHG 87
 QY 67 LVGGASNIWNLNPNNSLGFILADAGFDVVMGNSRGNAMSKKHTLSIDODEFWAFSYDEM 126
 DB 88 LLADSSNVTNLANSLSLGFILADAGFDVVMGNSRGNAMSKKHTLSVSODEFWAFSYDEM 147
 QY 127 ARFDLPVAVINFILQKTGOEKIYVYVGSQGTVMGFIASFSTMPPELAOKIKMYFALAPIATVK 186
 DB 148 AKYDLPAFASINFILNKTKGQEVYVGHSGQTTIGFIASFQIPELAKRIKMFALGPVASVA 207
 QY 187 HAKSPGTFKLLPDMNKGFLGKKEFLYQTRFLRQLVYLQGVILQICSNIMLLGGF 246
 DB 208 FCTSPMAKLGRLPDHLIKDLFGDKFELPQSAFLKWLGTHTVTHVILKELCGLLQCGF 267
 QY 247 NTNNMNSRASVVAHTLACTSVQNLHWSQAVNSGELRAFDWGSSTKLEKCNQPTVR 306
 DB 268 NERNLNMSRDVYTHSPAGTSVQNMHLWSQAVKFKQFAFDWGSSTKLEKCNQPTVR 327
 QY 307 YRVDMTVPTAMWTGGQDLNPNEDVYKLLSEVNTLIYHKNIPPEWAHVDFIWLGDAPHRM 366
 DB 328 YNVKMDLVPVAVNSGGHDLADYVDVNNILLTQITLVFHSIPEWEHLDFIWLGDAPWRL 387
 QY 367 YNEIHLMOQ 376
 DB 388 YNKIINLMRK 397

RESULT 10

AAAB66061
 ID AAB66061 standard; Protein; 399 AA.

AC AAB66061;

DT 30-MAR-2001 (first entry)

DE Human lysosomal acid lipase protein.

QY Gene therapy; TANGO protein; INTERCEPT protein; neurological disorder;
 KW central nervous system; focal brain disorder; bipolar affective disorder;
 KW global-diffuse cerebral disorder; cerebrovascular; Alzheimer's disease;
 KW senile dementia; Huntington's disease; amyotrophic lateral sclerosis;
 KW Parkinson's; Gilles de la Tourette's syndrome; hypertension; sleep;
 KW neuropsychiatric; psychoactive substance use; anxiety.

OS Homo sapiens.

XX WO2000077239-A2.

XX 21-DEC-2000.

XX 24-MAY-2000; 2000WO-US14858.

XX 14-JUN-1999; 99US-0333159.

XX (MILL-) MILLENNIUM PHARM INC.

XX McCarthy SA, Fraser CC, Sharp JD, Barnes TW;

XX WPI; 2001-032313/04.

XX TANGO and INTERCEPT nucleic acids, proteins, and antibodies, useful for
 PT screening assays and diagnostic assays and for the treatment of
 PT neurological diseases such as Alzheimer's, Parkinson's and Huntington's
 PT disease -

XX Claim 8; Fig 6; 359pp; English.

XX The present invention relates to TANGO or INTERCEPT proteins and coding
 CC sequences (see AAF45121-F45136 and AAF45138-F45139 and AAB66031-B66057,
 CC AAB66064-B66083 and AAB66085). The TANGO/INTERCEPT proteins and coding
 CC sequences are useful for the treatment of neurological disorders such as

CC central nervous system (CNS) disorders, CNS-related disorders, focal
 CC brain disorders, Global-diffuse cerebral disorders and other
 CC neurological and cerebrovascular disorders. The CNS disorders include
 CC Alzheimer's disease, senile dementia, Huntington's disease, amyotrophic
 CC lateral sclerosis, Parkinson's, Gilles de la Tourette's syndrome,
 CC autonomic function disorders such as hypertension and sleep disorders,
 CC neuropsychiatric disorders, psychoactive substance use disorders,
 CC anxiety, and bipolar affective disorder. The present sequence is a
 CC sequence used in a sequence homology alignment with the TANGO/INTERCEPT
 CC sequences of the present invention.

XX Sequence 399 AA;

QY Query Match 62.1%; Score 1289; DB 22; Length 399;
 DB Best Local Similarity 63.5%; Pred. No. 1.4e-123;
 DB Matches 235; Conservative 55; Mismatches 80; Indels 0; Gaps 0;

QY 7 AVDPFAFMTSEIIHQGYPCPEEYEVATEDGYILSVNRIPLGVLPKKTGSRPVVLLQHG 66
 DB 28 AVDPETNMVSEIIISYMGFPSEYLVETEDGYILCLNRIPLHGRKXSDKGPFPVFLQHG 87
 QY 67 LVGGASNIWNLNPNNSLGFILADAGFDVVMGNSRGNAMSKKHTLSIDODEFWAFSYDEM 126
 DB 88 LLADSSNVTNLANSLSLGFILADAGFDVVMGNSRGNAMSKKHTLSVSODEFWAFSYDEM 147
 QY 127 ARFDLPVAVINFILQKTGOEKIYVYVGSQGTVMGFIASFSTMPPELAOKIKMYFALAPIATVK 186
 DB 148 AKYDLPAFASINFILNKTKGQEVYVGHSGQTTIGFIASFQIPELAKRIKMFALGPVASVA 207
 QY 187 HAKSPGTFKLLPDMNKGFLGKKEFLYQTRFLRQLVYLQGVILQICSNIMLLGGF 246
 DB 208 FCTSPMAKLGRLPDHLIKDLFGDKFELPQSAFLKWLGTHTVTHVILKELCGLLQCGF 267
 QY 247 NTNNMNSRASVVAHTLACTSVQNLHWSQAVNSGELRAFDWGSSTKLEKCNQPTVR 306
 DB 268 NERNLNMSRDVYTHSPAGTSVQNMHLWSQAVKFKQFAFDWGSSTKLEKCNQPTVR 327
 QY 307 YRVDMTVPTAMWTGGQDLNPNEDVYKLLSEVNTLIYHKNIPPEWAHVDFIWLGDAPHRM 366
 DB 328 YNVKMDLVPVAVNSGGHDLADYVDVNNILLTQITLVFHSIPEWEHLDFIWLGDAPWRL 387
 QY 367 YNEIHLMOQ 376
 DB 388 YNKIINLMRK 397

RESULT 11

AAO18227
 ID AAO18227 standard; Protein; 392 AA.

XX AAO18227;

XX 18-SEP-2002 (first entry)

XX Human lysosomal acid lipase related protein.

XX Human; lysosomal acid lipase; chromosome 10; lipemic; Wolman disease;
 KW cholesteryl ester storage disease.

XX Homo sapiens.

XX US6387680-B1.

XX 14-MAY-2002.

XX 29-MAR-2001; 2001US-0820001.

XX 29-MAR-2001; 2001US-0820001.

XX (PEKE) PE CORP NY.

XX Merkulov GV, Ketchum KA, Di Francesco V, Beasley EM;

XX

XX 30-MAR-2001 (first entry)
 XX Human TANGO 294 extracellular domain.
 XX TANGO protein; INTERCEPT protein; neurological disorder;
 XX central nervous system; focal brain disorder; bipolar affective disorder;
 XX Global-diffuse cerebral disorder; cerebrovascular; Alzheimer's disease;
 XX senile dementia; Huntington's disease; amyotrophic lateral sclerosis;
 XX Parkinson's; Gilles de la Tourette's syndrome; hypertension; sleep;
 XX neuropsychiatric; psychoactive substance use; anxiety.
 OS Homo sapiens.
 XX W0200077239-A2.
 XX 21-DEC-2000.
 XX 24-MAY-2000; 2000WO-US14858.
 XX 14-JUN-1999; 99US-0333159.
 XX (MILL-) MILLENNIUM PHARM INC.
 XX McCarthy SA, Fraser CC, Sharp JD, Barnes TM;
 XX WPI; 2001-032313/04.
 XX TANGO and INTERCEPT nucleic acids, proteins, and antibodies, useful for
 XX screening assays and diagnostic assays and for the treatment of
 XX neurological diseases such as Alzheimer's, Parkinson's and Huntington's
 XX disease -
 XX Claim 8; Page 326; 359pp; English.
 XX The present invention relates to TANGO or INTERCEPT proteins and coding
 XX sequences (see AAF45121-F45136 and AAF45138-F45139 and AAB66031-B66057,
 XX AAB66064-B66083 and AAB66085). The TANGO/INTERCEPT proteins and coding
 XX sequences are useful for the treatment of neurological disorders such as
 XX central nervous system (CNS) disorders, CNS-related disorders, focal
 XX brain disorders, global-diffuse cerebral disorders and other
 XX neurological and cerebrovascular disorders. The CNS disorders include
 XX Alzheimer's disease, senile dementia, Huntington's disease, amyotrophic
 XX lateral sclerosis, Parkinson's, Gilles de la Tourette's syndrome,
 XX autonomic function disorders such as hypertension and sleep disorders,
 XX neuropsychiatric disorders, psychoactive substance use disorders,
 XX anxiety, and bipolar affective disorder.
 XX Sequence 221 AA;
 Query Match 56.1%; Score 1165; DB 22; Length 221;
 Best Local Similarity 100.0%; Pred. No. 3.1e-111;
 Matches 221; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 QY 1 VHMPTKAVDPAPFNNISEIIHQGQPCPEYEVEATEDGYLLSVNRIPRGLVQPKTKTSRPV 60
 DB 1 VHMPTKAVDPAPFNNISEIIHQGQPCPEYEVEATEDGYLLSVNRIPRGLVQPKTKTSRPV 60
 QY 61 VLLQGLVGGASNMISNLPNNSLGFTILADAGFDVWNGNSRGNWSRKHKTLSIDQDEFWA 120
 DB 61 VLLQGLVGGASNMISNLPNNSLGFTILADAGFDVWNGNSRGNWSRKHKTLSIDQDEFWA 120
 QY 121 PSYDEMAFEDLPVNFNFILOKTGQEKIYVGYSGTMTGFIAPSTMPELAQIKMYFALA 180
 DB 121 PSYDEMAFEDLPVNFNFILOKTGQEKIYVGYSGTMTGFIAPSTMPELAQIKMYFALA 180
 QY 181 PIATVGHAKSPGTGFTLLPDMNMIKGLFGKKEFLYQTRFLRQ 221
 DB 181 PIATVGHAKSPGTGFTLLPDMNMIKGLFGKKEFLYQTRFLRQ 221
 RESULT 14
 AAP50322

II AAP50322 standard; protein; 395 AA.
 XX AAP50322;
 AC 25-MAR-2003 (updated)
 DT 17-JAN-1992 (first entry)
 XX Rat lingual lipase protein.
 DE Lingual lipase; enzyme; EC-3.1.1.3; ss.
 KW Rattus rattus.
 OS GB2142337-A.
 XX 16-JAN-1985.
 PD 29-JUN-1984; 84GB-0016581.
 XX 29-JUN-1984; 84GB-0016581.
 XX (CLLT) CELLTECH LTD.
 XX (WILL/) WILLIAMSON R.
 XX Carey NH, Williamson R;
 XX WPI; 1985-014450/03.
 XX N-PSDB; AAN50385.
 XX New lingual lipase protein for treatment of lipase deficiency - also
 XX new pre-lingual lipase protein and related products
 XX Disclosure; Fig 5; 15pp; English.
 XX This protein may be expressed in a transformant host organism
 XX and may be used for the treatment of lipase deficiency.
 XX (Updated on 25-MAR-2003 to correct PR field.)
 XX (Updated on 25-MAR-2003 to correct PA field.)
 XX Sequence 395 AA;
 Query Match 55.5%; Score 1153; DB 6; Length 395;
 Best Local Similarity 57.5%; Pred. No. 1.3e-109;
 Matches 212; Conservative 63; Mismatches 94; Indels 0; Gaps 0;
 QY 9 DPEAFMNTSEIIHQGQPCPEYEVEATEDGYLLSVNRIPRGLVQPKTKTSRPVLLQGLV 68
 DB 27 NPEANMNTSQMTYWGYPQCEYEVEATEDGYLLSVNRIPRGLVQPKTKTSRPVLLQGLV 86
 QY 69 GGASNMISNLPNNSLGFTILADAGFDVWNGNSRGNWSRKHKTLSIDQDEFWAFSDYDEMA 128
 DB 87 ASATNWIANLPNNSLAFMLADAGYDVLGNSRGNWTSRKNVYSPDSVEFWAFSDYDEMA 146
 QY 129 FDLPAVINFILOKTGQEKIYVGYSGTMTGFIAPSTMPELAQIKMYFALAPIATVKA 188
 DB 147 VDLPATINFILOKTGQEKIYVGYSGTMTGFIAPSTMPELAQIKMYFALAPIATVKA 206
 QY 189 RSPGKFKLLPDMNMIKGLFGKKEFLYQTRFLRQVILVLCGOVILDLQCSNMLLGGFNT 248
 DB 207 QSLPKKISIFITFTFKLFGKKNFLPHYFDFFLGTEVCSEVLDLCSNMLLFGCFDX 266
 QY 249 NNMNMSRASVYAAHTLACTSVQNTILHMSQAVNSGELAFDWSGTSKLEKNCQPTPVYR 308
 DB 267 KNLNVSREDVYLGHNPAGTSVQDFLHWAQLVRSKGFQAFNWSGSPQNMLHYNQKTPPEYD 326
 QY 309 VEDMTVPAMWTGGQDWLSNEDVYKMLSEVYTNLIYHKNIPWAHVDPIGLDAPHMYN 368
 DB 327 VSAMTVPVAVWNGNDILADPDVAMLLPKLSNLLFHKELAYNHLDFIANDAPQEVYN 386
 QY 369 EIIHLMQOE 377
 DB 387 ENISMMAED 395

RESULT 15
AAR37302
ID AAR37302 standard; Protein; 398 AA.
XX AC AAR37302;
XX DT 25-MAR-2003 (updated)
XX DT 20-SEP-1993 (first entry)
XX XX
XX DE RGL precursor.
XX
XX Rabbit gastric lipase; RGL; pRGLM2.1; fat; bioconversion;
XX hydrolysis; transesterification.
XX Oryctolagus cuniculus.
XX
XX FH Key Location/Qualifiers
XX FT Protein 23..398
XX FT /note= "claim 1; page 10-11"
XX PN EP542629-A1.
XX PD 19-MAY-1993.
XX PF 12-NOV-1992; 92EP-0403055.
XX PR 13-NOV-1991; 91FR-0013948.
XX PA (LJOU) INST RECH JOUVEINAL.
XX PI Benicourt C, Blanchard C, Junien J;
XX WPI; 1993-161080/20.
XX DR N-PSDB; AAQ42310.
XX
XX Rabbit gastric lipase, its precursor and their DNA - useful for
XX treating conditions linked to gastric lipase deficiency, such as
XX mucoviscidiosis and pancreatic exocrine insufficiency
XX
XX Claim 1; Fig 7; 31pp; French.
XX
XX RGL, opt. used with other lipases, are useful therapeutically
XX (1) to facilitate absorption of ingested fats in patients deficient
XX in endogenous gastric lipase and (2) to treat disorders caused by
XX inadequate lipase prodn. esp. mucoviscidiosis or pancreatic exocrine
XX deficiency. It can also be used (partic. when immobilised) for
XX enzymatic bioconversion, e.g. hydrolysis or transesterification.
XX (Updated on 25-MAR-2003 to correct PN field.)
XX
XX SQ Sequence 398 AA;

Query Match 55.4%; Score 1150.5; DB 14; Length 398;
Best Local Similarity 56.8%; Pred. NO. 2.4e-109;
Matches 213; Conservative 65; Mismatches 94; Indels 3; Gaps 1;

QY 4 PTKAVDPFAFNIISIIHQGYPCBEYEVATEDGYILSVNRIPRGLVQPKTKGSRPVLL 63
DB 26 PT---NPEVNNNISQISWGYPSKYEIVTDEDGYILEVNRIPYKKNNGRQRPVFL 82
QY 64 QHGLVGGASNIWISLPPNNSLGFILADAGFDVMMGNSRGNWSRKHKTLSDQDBFWAFSY 123
DB 83 QHGLLASASNIWISLPPNNSLGFILADAGYVGLGNSRGNWTSRRNLVYSPDSVEFWAFSF 142
QY 124 DEMARFDLPVNFILQTKGKLYVGYSGTGTGTFTARSTMPETLQAKIMYFALAPIA 183
DB 143 DEMAKYDLPATIDFIVKGTGKLYVGYSGTGTGTFTARSTMPETLQAKIMYFALAPIA 202
QY 184 TVKHAKSPGTKFLLLPDMWIKGLFKKFLYQTRFLRLQVLYLQGVILQICSNIMLL 243
DB 203 TVKYTKSLVNLKRFIPPTMFKIIFGDKIFYPHNPFDFQATQVCSRETLNVICSNALFII 262
QY 244 GGFNTNNNMRSASYAAHTLAGTSVQNILHWSQAVNSGELRAPDWGSETKLEKCNQPT 303

Search completed: February 19, 2004, 14:20:38
Job time : 34.1402 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2004 Compugen Ltd.

OM protein - protein search, using sw model

Run on: February 19, 2004, 14:19:24 ; Search time 13.4317 Seconds
(without alignments)
1228.526 Million cell updates/sec

Title: US-10-042-431-49

Perfect score: 2076

Sequence: 1 VHMPTKAVDPEAFNMISEII.....IHLMQEETNLSQRCBAVL 390

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 328717 seqs, 42310858 residues

Total number of hits satisfying chosen parameters: 328717

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : Issued Patents AA.*

- 1: /cgm2_6/ptodata/2/iaa/5A.COMB.pep.*
- 2: /cgm2_6/ptodata/2/iaa/5B.COMB.pep.*
- 3: /cgm2_6/ptodata/2/iaa/6A.COMB.pep.*
- 4: /cgm2_6/ptodata/2/iaa/6B.COMB.pep.*
- 5: /cgm2_6/ptodata/2/iaa/PTCUS.COMB.pep.*
- 6: /cgm2_6/ptodata/2/iaa/backfiles1.pep.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	1282	61.8	392	4	US-09-820-001-4
2	1161	55.9	377	1	US-08-227-108-17
3	1161	55.9	377	2	US-09-073-674-17
4	1150.5	55.4	379	2	US-08-227-108-18
5	1150.5	55.4	379	2	US-09-073-674-18
6	1141.5	55.0	375	4	US-09-348-930A-6
7	1141.5	55.0	379	1	US-08-227-108-3
8	1141.5	55.0	379	2	US-09-073-674-3
9	1141.5	55.0	379	4	US-09-348-930A-2
10	1141.5	55.0	380	1	US-08-227-108-5
11	1141.5	55.0	380	2	US-09-073-674-5
12	1127	54.3	378	4	US-09-186-489-2
13	1127	54.3	378	4	US-10-043-665B-2
14	1109	53.4	398	4	US-09-348-930A-9
15	1098.5	52.9	380	1	US-08-227-108-16
16	1098.5	52.9	380	2	US-09-073-674-16
17	1023	49.3	395	4	US-09-820-001-2
18	1003	48.3	325	4	US-09-348-930A-4
19	143.5	6.9	340	4	US-09-252-991A-30091
20	105	5.1	430	1	US-08-484-105-22
21	105	5.1	430	1	US-08-484-106-22
22	103	5.0	292	4	US-09-328-352-7538
23	98.5	4.7	388	1	US-08-232-519-2
24	98.5	4.7	388	1	US-08-456-956-2
25	96.5	4.6	287	4	US-09-584-568C-4
26	96.5	4.6	1752	4	US-09-865-621A-2
27	95	4.6	346	2	US-08-602-359A-34

28	91	4.4	298	4	US-09-355-166-6	Sequence 6, Appli
29	90	4.3	277	4	US-09-424-349A-6	Sequence 6, Appli
30	89	4.3	312	4	US-09-107-532A-5684	Sequence 5684, Ap
31	89	4.3	494	4	US-09-134-001C-4475	Sequence 4475, Ap
32	89	4.3	684	4	US-09-252-991A-18416	Sequence 18416, A
33	89	4.3	1674	2	US-08-968-542C-12	Sequence 12, Appl
34	88	4.2	826	4	US-09-328-352-7515	Sequence 7515, Ap
35	87.5	4.2	297	4	US-09-584-568C-7	Sequence 7, Appli
36	87	4.2	270	4	US-09-107-532A-4778	Sequence 4778, Ap
37	87	4.2	624	2	US-08-756-317-9	Sequence 9, Appli
38	86.5	4.2	685	4	US-08-671-757A-12	Sequence 12, Appl
39	86.5	4.2	869	4	US-08-483-101-15	Sequence 15, Appl
40	84.5	4.1	173	4	US-09-328-352-8078	Sequence 8078, Ap
41	84.5	4.1	279	4	US-09-328-352-5958	Sequence 5958, Ap
42	84.5	4.1	338	2	US-08-602-359A-40	Sequence 40, Appl
43	84	4.0	347	4	US-09-328-352-5921	Sequence 5921, Ap
44	84	4.0	719	3	US-09-386-607-2	Sequence 2, Appli
45	84	4.0	719	4	US-09-645-707B-2	Sequence 2, Appli

ALIGNMENTS

RESULT 1

US-09-820-001-4
; Sequence 4, Application US/09820001
; Patent No. 6387680
; GENERAL INFORMATION:
; APPLICANT: MERKULOV, Gennady et al
; TITLE OF INVENTION: ISOLATED HUMAN LIPASE PROTEINS, NUCLEIC
; TITLE OF INVENTION: ACID MOLECULES ENCODING HUMAN LIPASE PROTEINS, AND USES
; TITLE OF INVENTION: THEREOF
; FILE REFERENCE: CL001186
; CURRENT APPLICATION NUMBER: US/09/820,001
; CURRENT FILING DATE: 2001-03-29
; NUMBER OF SEQ ID NOS: 4
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 4
; LENGTH: 392
; TYPE: PRT
; ORGANISM: Human
US-09-820-001-4

Qy	7	AVDPEAFNMISEII	CHQGYPCSEYEVATEDGVILSVNRIPLGLVQPKTGRPVVLLQHG	66
Db	26	AVDPTNMNVSEII	SYWGPSEIYVETEDGVILCNRIPIGRKNHSKSGKPKVFLQHG	85
Qy	67	LVGGASNWTISNLPNNSLGF	ILADAGFDVVMGNSRGNAMSRKHKTLSDIDQDEFWAFSYDEM	126
Db	86	LLADSSNVTNLANSLSGFI	LADAGFDVVMGNSRGNAMSRKHKTLSDVSDQDEFWAFSYDEM	145
Qy	127	AFDIPAVINFILOKTQGEKI	YVGYSGTTFGTFAPSTMPPELAQKIMYFALAPIATVK	186
Db	146	AKYDLPASINFILKTKGQEV	YVGHSGTTFGTFAPSTMPPELAQKIMYFALAPIATVK	205
Qy	187	HAKSPGTEKLLPDMWIKGL	FGKGEFLYQTRFLQVLYLQGVILDOQICSNIMLLCGF	246
Db	206	FCTSPMAKLGRLPDHLIKD	FGDKKEFLPQSAFLKWLGVTHVILKELCGNLCLCGF	265
Qy	247	NTNNMNSASVYAHTLAGTS	VQNLHWSQVNSGELRAPDWSGSETKNLEKCNQTPVR	306
Db	266	NEENLNMNSVDVYTHSPAGT	SVQNLHWSQVNSGELRAPDWSGSETKNLEKCNQTPVR	325
Qy	307	YVRDNTVPTAMWTGQDMLSN	PEDVCMLLSEVTLNLIYHKNIPEWAHVDPIWGLDAPHRM	366
Db	326	YVCKMLVPTAVWSGGHDLAD	VDVYNILLTQITNLVPHESIPWEHLDPIWGLDAPWEL	385
Qy	367	YNEIHL	373	

Query Match 61.8%; Score 1282; DB 4; Length 392;
Best Local Similarity 63.8%; Pred. No. 3.1e-129;
Matches 234; Conservative 53; Mismatches 80; Indels 0; Gaps 0;

QY 309 VRDVTPTAMTGGQDWSLNPEDYKMLSEVTNLIYHKNIPWAHVDIFWGLDAPHRMYN 368
 Db 309 VSAMTVPVAVWNGNDIILADPDQVAMLLFKLSNLLFHKELIAYNLHDFIWMADAPQEVYN 368
 QY 369 EIIHLMQOE 377
 Db 369 EMISMMAED 377

RESULT 4

US-08-227-108-18
 ; Sequence 18, Application US/08227108
 ; Patent No. 5807726
 ; GENERAL INFORMATION:
 ; APPLICANT: Blanchard, Claire
 ; APPLICANT: Benicourt, Claude
 ; APPLICANT: Junien, Jean-Louis
 ; TITLE OF INVENTION: Recombinant Dog Gastric Lipase
 ; NUMBER OF SEQUENCES: 21
 ; CORRESPONDENCE ADDRESS:
 ; ADDRESSEE: Pennie & Edmonds
 ; STREET: 1155 Avenue of the Americas
 ; CITY: New York
 ; STATE: New York
 ; COUNTRY: U.S.A.
 ; ZIP: 10036

COMPUTER READABLE FORM:
 ; MEDIUM TYPE: Floppy disk
 ; COMPUTER: IBM PC compatible
 ; OPERATING SYSTEM: PC-DOS/MS-DOS
 ; SOFTWARE: Patentin Release #1.0, Version #1.25
 ; CURRENT APPLICATION DATA:
 ; APPLICATION NUMBER: US/08/227.108
 ; FILING DATE: 03-APR-1994
 ; CLASSIFICATION: 435

ATTORNEY/AGENT INFORMATION:
 ; NAME: Panucci, Allan A.
 ; REGISTRATION NUMBER: 30,256
 ; REFERENCE/DOCKET NUMBER: 7620-033

TELECOMMUNICATION INFORMATION:
 ; TELEPHONE: 212 790-9090
 ; TELEFAX: 212 869-8864/9741
 ; TELEX: 66141 PENNIE

INFORMATION FOR SEQ ID NO: 18:
 ; SEQUENCE CHARACTERISTICS:
 ; LENGTH: 379 amino acids
 ; TYPE: amino acid
 ; STRANDEDNESS: single
 ; TOPOLOGY: linear
 ; MOLECULE TYPE: protein
 ; US-08-227-108-18

Query Match 55.4%; Score 1150.5; DB 1; Length 379;
 Best Local Similarity 56.8%; Pred. No. 4.1e-115;
 Matches 213; Conservative 65; Mismatches 94; Indels 3; Gaps 1;

QY 4 PTKAVDPEAFMNISEIIHQHGPCYEYVATEDGVILSVNRIPLGVQPKTGRSPVLL 63
 Db 7 PT---NPEVNMNISQISWGYPSKEYEVVTEGIVLEVNRIPIYKQKNSGNRGQRPVFL 63
 QY 64 QHGLVGASNWTISNLPNNISLGFILADAGFVWNGSRGNWSRKHKTLSDIDQDEFWAFSY 123
 Db 64 QHGLLASASNWTISNLPNNISLGFILADAGFVWNGSRGNWTNRRNLYYSPDSVEFWAFSP 123
 QY 124 DEMARFDLPVINFILQKTQGEKIYVYVGSQGTTFGTAFTSTMPELAOKIKMYFALAPIA 183
 Db 124 DEMAKYDLPATIDFIVKGTQGEKLYVGHVSQGTTFGTAFTSTNPKLAERIKTFYALAPVA 183
 QY 184 TVKHAKSPGTEKFLLLPDMNKGIFGKGFELYQTRFLRQLVILCGQVILDOICSNIMLLL 243
 Db 184 TVKYTKSLVNLKRFIPPTMFKIIFGDKIPYPHNFDPQFLATQVCSRETNLVICSNALFII 243
 QY 244 GGFNTNNMWSRASVYAAHTLAGTSVQNLHWSQVNSGELRAFDWNGSETKXLEKCNQPT 303

Db 244 CGFDSANLNSRLDVTVSHNPAGTSVQNLHWTQVKGSGNPOAFNWSGPAQNVWFHNPQT 303
 QY 304 PVRVRYRDMTPTAMTGGQDWSLNPEDYKMLSEVTNLIYHKNIPWAHVDIFWGLDAP 363
 Db 304 PPTVNTAMNVEIAVWSGNDWLADPDQVDLKLKLSNLIYHKEILPYNLHDFIWMANAP 363
 QY 364 HMYNEIILHMQOE 378
 Db 364 QEVINEIISMAEDK 378

RESULT 5

US-09-073-674-18
 ; Sequence 18, Application US/09073674
 ; Patent No. 5998189

GENERAL INFORMATION:
 ; APPLICANT: Blanchard, Claire
 ; APPLICANT: Benicourt, Claude
 ; APPLICANT: Junien, Jean-Louis
 ; TITLE OF INVENTION: Recombinant Dog Gastric Lipase
 ; NUMBER OF SEQUENCES: 21
 ; CORRESPONDENCE ADDRESS:
 ; ADDRESSEE: Warner-Lambert Company
 ; STREET: 2800 Plymouth Road
 ; CITY: Ann Arbor
 ; STATE: Michigan
 ; COUNTRY: U.S.A.
 ; ZIP: 48105

COMPUTER READABLE FORM:
 ; MEDIUM TYPE: Floppy disk
 ; COMPUTER: IBM PC compatible
 ; OPERATING SYSTEM: PC-DOS/MS-DOS
 ; SOFTWARE: Patentin Release #1.0, Version #1.25
 ; CURRENT APPLICATION DATA:
 ; APPLICATION NUMBER: US/09/073,674
 ; FILING DATE:

CLASSIFICATION:
 ; ATTORNEY/AGENT INFORMATION:
 ; NAME: Criesey, Todd M.
 ; REGISTRATION NUMBER: 37,807
 ; REFERENCE/DOCKET NUMBER: 5072-DI-66-TWC
 ; TELECOMMUNICATION INFORMATION:
 ; TELEPHONE: 734 622-7530
 ; TELEFAX: 734 622-1553
 ; TELEX:

INFORMATION FOR SEQ ID NO: 18:
 ; SEQUENCE CHARACTERISTICS:
 ; LENGTH: 379 amino acids
 ; TYPE: amino acid
 ; STRANDEDNESS: single
 ; TOPOLOGY: linear
 ; MOLECULE TYPE: protein
 ; US-09-073-674-18

Query Match 55.4%; Score 1150.5; DB 2; Length 379;
 Best Local Similarity 56.8%; Pred. No. 4.1e-115;
 Matches 213; Conservative 65; Mismatches 94; Indels 3; Gaps 1;

QY 4 PTKAVDPEAFMNISEIIHQHGPCYEYVATEDGVILSVNRIPLGVQPKTGRSPVLL 63
 Db 7 PT---NPEVNMNISQISWGYPSKEYEVVTEGIVLEVNRIPIYKQKNSGNRGQRPVFL 63
 QY 64 QHGLVGASNWTISNLPNNISLGFILADAGFVWNGSRGNWSRKHKTLSDIDQDEFWAFSY 123
 Db 64 QHGLLASASNWTISNLPNNISLGFILADAGFVWNGSRGNWTNRRNLYYSPDSVEFWAFSP 123
 QY 124 DEMARFDLPVINFILQKTQGEKIYVYVGSQGTTFGTAFTSTMPELAOKIKMYFALAPIA 183
 Db 124 DEMAKYDLPATIDFIVKGTQGEKLYVGHVSQGTTFGTAFTSTNPKLAERIKTFYALAPVA 183
 QY 184 TVKHAKSPGTEKFLLLPDMNKGIFGKGFELYQTRFLRQLVILCGQVILDOICSNIMLLL 243

Db 184 TVKTKSLVNLKRLPIPTMTFKIIFGDKIFYPHNFDDQFLATQVCSRETNLVICSNALFTI 243

Qy 244 GGFNTNNMNSRASVYAAHTLAGTSVQNLHWSQAVNSGELAFDWSGSETKLEKCNQPT 303

Db 244 CGFDSANLNSRLDVTYVSHNPAGTSVQNLHWSQAVNSGELAFDWSGSETKLEKCNQPT 303

Qy 304 PVRVVRDMVTPTAMTGGODWLSNPDVVKMLSEVTNLIYHKNIPKIAPEWAFVDFIWLGLDAP 363

Db 304 PPTVNTVAMNVPPIAVWSGGNDLADPDQVLLPKULNLIYHKNIPKIAPEWAFVDFIWLGLDAP 363

Qy 364 HRMYNEIHLMOQEE 378

Db 364 QEVNIEIISMAXDK 378

RESULT 6

US-09-348-930A-6

Sequence 6, Application US/09348930A

Patent No. 6573431

GENERAL INFORMATION:

APPLICANT: Lenee, P

APPLICANT: Gruber, V

APPLICANT: Baudino, S

APPLICANT: Merot, B

APPLICANT: Benicourt, C

APPLICANT: Cudrey, C

TITLE OF INVENTION: Recombinant Pseudoal Lipases and Polypeptides Derivatives Prod

TITLE OF INVENTION: Plants, Processes for Obtaining Them and Their Uses

FILE REFERENCE: 18433/2012

CURRENT APPLICATION NUMBER: US/09/348,930A

CURRENT FILING DATE: 1999-07-02

PRIOR APPLICATION NUMBER: 08/945,321

PRIOR FILING DATE: 1998-02-12

PRIOR APPLICATION NUMBER: PCT/FR96/00606

PRIOR FILING DATE: 1996-03-19

NUMBER OF SEQ ID NOS: 37

SOFTWARE: Patent in version 3.0

SEQ ID NO 6

LENGTH: 375

TYPE: PRT

ORGANISM: Canis familiaris

US-09-348-930A-6

Query Match 55.0%; Score 1141.5; DB 4; Length 375;

Best Local Similarity 56.5%; Pred. No. 3.7e-114;

Mismatches 210; Conservative 67; Mismatches 90; Indels 5; Gaps 3;

Qy 4 PTKAVDPEAFNISEIIOHQGYPCVEEVEVATEDGYILSVNRIPLGLVQPKTKGSRPVLL 63

Db 3 PT---NPEVTNLSQMTYNGYPAEEVEVATEDGYILGIDRIPYGRKNSENIGRRPVAF 59

Qy 64 QHGLVGGASNWSISLNNPNSLGFILADAGFDVWNGNSRGNWSRKHKTLSIDQDEFWAF 123

Db 60 QHGLASATNWSISLNNPNSLGFILADAGFDVWNGNSRGNWSRKHKTLSIDQDEFWAF 119

Qy 124 DEMARFDLPVINEITLOKTGOEKIYGYSGTMTGFIAPSTWPELAOKIKMYFALAPIA 183

Db 120 DEMAKYDLPATIDFILKKTGQDKLHYVHSGTITGFIAPSTWPELAOKIKMYFALAP 179

Qy 184 TVKHAQSPGTEKLLPDMNMIKGLFKKEFLYQTRFLRQ-LVYLCGVQVILDCSNIMLL 242

Db 180 TVKYTETLLNKLMLVPSFLPKLIFGNKIF-YPHFFDQFLATEVCSRETVDLLCSNALFI 238

Qy 243 LGGFTNNMNSRASVYAAHTLAGTSVQNLHWSQAVNSGELAFDWSGSETKLEKCNQPT 302

Db 239 ICGFDTNMLNLSRLDVTYVSHNPAGTSVQNLHWSQAVNSGELAFDWSGSETKLEKCNQPT 298

Qy 303 TPVRYVRDMVTPTAMTGGODWLSNPDVVKMLSEVTNLIYHKNIPKIAPEWAFVDFIWLGLD 362

Db 299 MPYYNLTDMHVPPIAVWSGGNDLADPDQVLLPKULNLIYHKNIPKIAPEWAFVDFIWLGLD 358

Qy 363 PHRYNEIHLMO 374

Db 359 QEVNIEIISMAXDK 378

RESULT 7

US-08-227-108-3

Sequence 3, Application US/08227108

Patent No. 5807726

GENERAL INFORMATION:

APPLICANT: Blanchard, Claire

APPLICANT: Benicourt, Claude

APPLICANT: Junien, Jean-Louis

TITLE OF INVENTION: Recombinant Dog Gastric Lipase

NUMBER OF SEQUENCES: 21

CORRESPONDENCE ADDRESS:

ADDRESSEE: Pennie & Edmonds

STREET: 1155 Avenue of the Americas

CITY: New York

STATE: New York

COUNTRY: U.S.A.

ZIP: 10036

COMPUTER READABLE FORM:

MEDIUM TYPE: Floppy disk

COMPUTER: IBM PC compatible

OPERATING SYSTEM: PC-DOS/MS-DOS

SOFTWARE: Patent in Release #1.0, Version #1.25

CURRENT APPLICATION DATA:

APPLICATION NUMBER: US/08/227,108

FILING DATE: 03-APR-1994

CLASSIFICATION: 435

ATTORNEY/AGENT INFORMATION:

NAME: Fanucci, Allan A.

REGISTRATION NUMBER: 30,256

REFERENCE/DOCKET NUMBER: 7620-033

TELECOMMUNICATION INFORMATION:

TELEPHONE: 212 790-9090

TELEFAX: 212 869-8864/9741

TELEX: 66141 PENNIE

INFORMATION FOR SEQ ID NO: 3:

SEQUENCE CHARACTERISTICS:

LENGTH: 379 amino acids

TYPE: amino acid

TOPOLOGY: linear

MOLECULE TYPE: protein

US-08-227-108-3

Query Match 55.0%; Score 1141.5; DB 1; Length 379;

Best Local Similarity 56.5%; Pred. No. 3.8e-114;

Mismatches 210; Conservative 67; Mismatches 90; Indels 5; Gaps 3;

Qy 4 PTKAVDPEAFNISEIIOHQGYPCVEEVEVATEDGYILSVNRIPLGLVQPKTKGSRPVLL 63

Db 7 PT---NPEVTNLSQMTYNGYPAEEVEVATEDGYILGIDRIPYGRKNSENIGRRPVAF 63

Qy 64 QHGLVGGASNWSISLNNPNSLGFILADAGFDVWNGNSRGNWSRKHKTLSIDQDEFWAF 123

Db 64 QHGLASATNWSISLNNPNSLGFILADAGFDVWNGNSRGNWSRKHKTLSIDQDEFWAF 123

Qy 124 DEMARFDLPVINEITLOKTGOEKIYGYSGTMTGFIAPSTWPELAOKIKMYFALAPIA 183

Db 124 DEMAKYDLPATIDFILKKTGQDKLHYVHSGTITGFIAPSTWPELAOKIKMYFALAP 183

Qy 184 TVKHAQSPGTEKLLPDMNMIKGLFKKEFLYQTRFLRQ-LVYLCGVQVILDCSNIMLL 242

Db 184 TVKYTETLLNKLMLVPSFLPKLIFGNKIF-YPHFFDQFLATEVCSRETVDLLCSNALFI 242

Qy 243 LGGFTNNMNSRASVYAAHTLAGTSVQNLHWSQAVNSGELAFDWSGSETKLEKCNQPT 302

Db 243 ICGFDTNMLNLSRLDVTYVSHNPAGTSVQNLHWSQAVNSGELAFDWSGSETKLEKCNQPT 302

Qy 303 TPVRYVRDMVTPTAMTGGODWLSNPDVVKMLSEVTNLIYHKNIPKIAPEWAFVDFIWLGLD 362

Db 303 MPYYNLTDMHVPPIAVWSGGNDLADPDQVLLPKULNLIYHKNIPKIAPEWAFVDFIWLGLD 362

QY 363 PHRYNEIHLHM 374
 Db 363 POAVNEIVSM 374

RESULT 8
 US-09-073-674-3
 ; Sequence 3, Application US/09073674
 ; Patent No. 598189
 ; GENERAL INFORMATION:
 ; APPLICANT: Blanchard, Claire
 ; APPLICANT: Benicourt, Claude
 ; APPLICANT: Junien, Jean-Louis
 ; TITLE OF INVENTION: Recombinant Dog Gastric Lipase
 ; NUMBER OF SEQUENCES: 21
 ; CORRESPONDENCE ADDRESS:
 ; ADDRESSEE: Warner-Lambert Company
 ; STREET: 2800 Plymouth Road
 ; CITY: Ann Arbor
 ; STATE: Michigan
 ; COUNTRY: U.S.A.
 ; ZIP: 48105
 ; COMPUTER READABLE FORM:
 ; MEDIUM TYPE: Floppy disk
 ; COMPUTER: IBM PC compatible
 ; OPERATING SYSTEM: PC-DOS/MS-DOS
 ; SOFTWARE: Patent in Release #1.0, Version #1.25
 ; CURRENT APPLICATION DATA:
 ; APPLICATION NUMBER: US/09/073,674
 ; FILING DATE:
 ; CLASSIFICATION:
 ; ATTORNEY/AGENT INFORMATION:
 ; NAME: Crissey, Todd M.
 ; REGISTRATION NUMBER: 37,807
 ; REFERENCE/DOCKET NUMBER: 5072-D1-66-TWC
 ; TELECOMMUNICATION INFORMATION:
 ; TELEPHONE: 734 622-7530
 ; TELEFAX: 734 622-1553
 ; TELEX:
 ; INFORMATION FOR SEQ ID NO: 3:
 ; SEQUENCE CHARACTERISTICS:
 ; LENGTH: 379 amino acids
 ; TYPE: amino acid
 ; TOPOLOGY: linear
 ; MOLECULE TYPE: protein
 ; US-09-073-674-3

Query Match 55.0%; Score 1141.5; DB 2; Length 379;
 Best Local Similarity 56.5%; Pred. No. 3.8e-114;
 Matches 210; Conservative 67; Mismatches 90; Indels 5; Gaps 3;

QY 4 PTKAVDPEAFNISEIIHQGYPCPEEYEVATEDGYILSVNRIPRGLVQPKTGRPVLL 63
 Db 7 PT---NPEVTWNISQMITWGYPAEYEVATEDGYILGIDRIPYGRKNSENIGRRPVAF 63
 QY 64 QHGLVGASNWSNLNPNNSLAFILADAGYDVLNWSRGNSRGNWRSKHTLSIDQDFWAFSY 123
 Db 64 QHGLLASATNWSNLNPNNSLAFILADAGYDVLNWSRGNSRGNWRSKHTLSIDQDFWAFSF 123
 QY 124 DEMARFDPFPAVINFILOKTGQEKIYVGYSGTMTMGFTAFSTMPBLAQIKMYPALAPIA 183
 Db 124 DEMAKYDLPATIDFILKKTGQDKLHYVGHSGQTIGFTAFSTNPKAKRIKTFYALAPVA 183
 QY 184 TVKHAKSPGTKEFLLLPDMNKGKFLYQTRFLRQ-LVYLGGQVILDOICSNIMLL 242
 Db 184 TVKYTEILLNKLMLVPFSLFKLIFGNKIF-YPHFFDQPLATEVCSRETVDLLCSNALFI 242
 QY 243 LGGFTNNMNSRASVYAAHTLAGTSVQNLHWSQAVNSGELRAPDWSGSETKNLEKNQ 302
 Db 243 ICGFDTNMLNSRLDVLISHNPAGTSVQNLHWSQAVNSGELRAPDWSGSETKNLEKNQ 302
 QY 303 TPVRYVRDVTPTAMWTGGQDWLSNPEDVKMLSEVTNLIYHKNIPEWAHVDFIWLDA 362
 Db 303 MPPYTNLTDMHVPVAVNNGNDLLADPHDVLNLSKLPNLIYHKKIPPNHLDFTWAMDA 362
 QY 363 PHRYNEIHLHM 374
 Db 363 POAVNEIVSM 374

RESULT 10
 US-08-227-108-5
 ; Sequence 5, Application US/08227108
 ; Patent No. 5807726
 ; GENERAL INFORMATION:

Db 303 MPPYTNLTDMHVPVAVNNGNDLLADPHDVLNLSKLPNLIYHKKIPPNHLDFTWAMDA 362
 QY 363 PHRYNEIHLHM 374
 Db 363 POAVNEIVSM 374

RESULT 9
 US-09-348-930A-2
 ; Sequence 2, Application US/09348930A
 ; Patent No. 6573431
 ; GENERAL INFORMATION:
 ; APPLICANT: Lenee, P
 ; APPLICANT: Gruber, V
 ; APPLICANT: Baudino, S
 ; APPLICANT: Merot, B
 ; APPLICANT: Benicourt, C
 ; APPLICANT: Cudrey, C
 ; TITLE OF INVENTION: Recombinant Predodenal Lipases and Polypeptides Derivatives Produ
 ; FILE REFERENCE: 18433/2012
 ; CURRENT APPLICATION NUMBER: US/09/348,930A
 ; CURRENT FILING DATE: 1999-07-02
 ; PRIOR APPLICATION NUMBER: 08/945,321
 ; PRIOR FILING DATE: 1998-02-12
 ; PRIOR APPLICATION NUMBER: PCT/FR96/00606
 ; PRIOR FILING DATE: 1996-03-19
 ; NUMBER OF SEQ ID NOS: 37
 ; SOFTWARE: Patent in version 3.0
 ; SEQ ID NO 2
 ; LENGTH: 379
 ; TYPE: PRT
 ; ORGANISM: Canis familiaris
 ; US-09-348-930A-2

Query Match 55.0%; Score 1141.5; DB 4; Length 379;
 Best Local Similarity 56.5%; Pred. No. 3.8e-114;
 Matches 210; Conservative 67; Mismatches 90; Indels 5; Gaps 3;

QY 4 PTKAVDPEAFNISEIIHQGYPCPEEYEVATEDGYILSVNRIPRGLVQPKTGRPVLL 63
 Db 7 PT---NPEVTWNISQMITWGYPAEYEVATEDGYILGIDRIPYGRKNSENIGRRPVAF 63
 QY 64 QHGLVGASNWSNLNPNNSLAFILADAGYDVLNWSRGNSRGNWRSKHTLSIDQDFWAFSY 123
 Db 64 QHGLLASATNWSNLNPNNSLAFILADAGYDVLNWSRGNSRGNWRSKHTLSIDQDFWAFSF 123
 QY 124 DEMARFDPFPAVINFILOKTGQEKIYVGYSGTMTMGFTAFSTMPBLAQIKMYPALAPIA 183
 Db 124 DEMAKYDLPATIDFILKKTGQDKLHYVGHSGQTIGFTAFSTNPKAKRIKTFYALAPVA 183
 QY 184 TVKHAKSPGTKEFLLLPDMNKGKFLYQTRFLRQ-LVYLGGQVILDOICSNIMLL 242
 Db 184 TVKYTEILLNKLMLVPFSLFKLIFGNKIF-YPHFFDQPLATEVCSRETVDLLCSNALFI 242
 QY 243 LGGFTNNMNSRASVYAAHTLAGTSVQNLHWSQAVNSGELRAPDWSGSETKNLEKNQ 302
 Db 243 ICGFDTNMLNSRLDVLISHNPAGTSVQNLHWSQAVNSGELRAPDWSGSETKNLEKNQ 302
 QY 303 TPVRYVRDVTPTAMWTGGQDWLSNPEDVKMLSEVTNLIYHKNIPEWAHVDFIWLDA 362
 Db 303 MPPYTNLTDMHVPVAVNNGNDLLADPHDVLNLSKLPNLIYHKKIPPNHLDFTWAMDA 362
 QY 363 PHRYNEIHLHM 374
 Db 363 POAVNEIVSM 374

RESULT 10
 US-08-227-108-5
 ; Sequence 5, Application US/08227108
 ; Patent No. 5807726
 ; GENERAL INFORMATION:

APPLICANT: Blanchard, Claire
APPLICANT: Benicourt, Claude
APPLICANT: Junien, Jean-Louis
TITLE OF INVENTION: Recombinant Dog Gastric Lipase
NUMBER OF SEQUENCES: 21
CORRESPONDENCE ADDRESS:
ADDRESSEE: Pennie & Edmonds
STREET: 1155 Avenue of the Americas
CITY: New York
STATE: New York
COUNTRY: U.S.A.
ZIP: 10036
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patent in Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/227,108
FILING DATE: 03-APR-1994
CLASSIFICATION: 435
ATTORNEY/AGENT INFORMATION:
NAME: Fanucci, Allan A.
REGISTRATION NUMBER: 30,256
REFERENCE/DOCKET NUMBER: 7620-033
TELECOMMUNICATION INFORMATION:
TELEPHONE: 212 790-9090
TELEFAX: 212 869-8864/9741
TELEX: 66141 PENNIE
INFORMATION FOR SEQ ID NO: 5:
SEQUENCE CHARACTERISTICS:
LENGTH: 380 amino acids
TYPE: amino acid
TOPOLOGY: linear
MOLECULE TYPE: protein
US-08-227-108-5

Query Match 55.0%; Score 1141.5; DB 1; Length 380;

Best Local Similarity 56.5%; Pred. No. 3.8e-114;
Matches 210; Conservative 67; Mismatches 90; Indels 5; Gaps 3;

QY 4 PTKAVDPEAFMNISEIIHQGVPCPEEYEVATEDGYILSVNRIIPRGLVQPKTGRPVVLL 63
DB 8 PT---NPEVTMNI-SQMTYGYPAEYEVVTEGYILGIDRIPYGRKNSENIGRRPVAF 64
QY 64 QHGLVGASNIWISNLPNSLGFILADAGFDVWNGSRGNANRKHKTLSIDODEFWAFSY 123
DB 65 QHGLLASATNIWISNLPNSLAFILADAGFDVWNGSRGNTWARRNLYSPDSVEFWAFSF 124
QY 124 DEMARFDPAVINFILOKTGOEKIYVYVGSQGTMGFIAPFTMPPELAOKIKMYFALAPIA 183
DB 125 DEMAKYDLPATIDFILKKTQDKLHVYVHSGQTTIGFIAPFTNPKLAKIKTFYALAPVA 184
QY 184 TVKHAKSPGKTFLLPDMMIKGLFGKKEFLYQTRFLRQ-LVIYLCQGVILDOICSNIMLL 242
DB 185 TVKYTETLLNKLMLVPSFLFKLIFGNKIF-YPHHFFDQFLATEVCSRETVDLLCSNALFI 243
QY 243 LGGFTNNMMSRASVYAAHTLAGTSVQNIILHWSQAVNSGELRAPDWGSETKNLEKCNOP 302
DB 244 ICGFTDNLNMSRLDVLNHPASTSVQNVLHWSQAVKSGKQADPWGSPVQNMHYSQS 303
QY 303 TPVRYVRDMVTPTAMWTGGQDWLSNPEDYKMLLSEVTNLIYHKNIPEWAHVDFIWGLDA 362
DB 304 MPYYNLTDMHVPYIAVWNGGNDLLADPHDVLNLSKLPNLIYHRKIPPNYHLDFIWAMDA 363
QY 363 PHEMYNEIHLM 374
DB 364 PQAVYNEIVSM 375

RESULT 11

US-09-073-674-5

; Sequence 5, Application US/09073674

Patent No. 5998189
GENERAL INFORMATION:
APPLICANT: Blanchard, Claire
APPLICANT: Benicourt, Claude
APPLICANT: Junien, Jean-Louis
TITLE OF INVENTION: Recombinant Dog Gastric Lipase
NUMBER OF SEQUENCES: 21
CORRESPONDENCE ADDRESS:
ADDRESSEE: Warner-Lambert Company
STREET: 2800 Plymouth Road
CITY: Ann Arbor
STATE: Michigan
COUNTRY: U.S.A.
ZIP: 48105
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patent in Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/09/073,674
FILING DATE:
CLASSIFICATION:
ATTORNEY/AGENT INFORMATION:
NAME: Crissey, Todd M.
REGISTRATION NUMBER: 37,807
REFERENCE/DOCKET NUMBER: 5072-D1-66-TMC
TELECOMMUNICATION INFORMATION:
TELEPHONE: 734 622-7530
TELEFAX: 734 622-1553
TELEX:
INFORMATION FOR SEQ ID NO: 5:
SEQUENCE CHARACTERISTICS:
LENGTH: 380 amino acids
TYPE: amino acid
TOPOLOGY: linear
MOLECULE TYPE: protein
US-09-073-674-5

Query Match 55.0%; Score 1141.5; DB 2; Length 380;

Best Local Similarity 56.5%; Pred. No. 3.8e-114;
Matches 210; Conservative 67; Mismatches 90; Indels 5; Gaps 3;

QY 4 PTKAVDPEAFMNISEIIHQGVPCPEEYEVATEDGYILSVNRIIPRGLVQPKTGRPVVLL 63
DB 8 PT---NPEVTMNI-SQMTYGYPAEYEVVTEGYILGIDRIPYGRKNSENIGRRPVAF 64
QY 64 QHGLVGASNIWISNLPNSLGFILADAGFDVWNGSRGNANRKHKTLSIDODEFWAFSY 123
DB 65 QHGLLASATNIWISNLPNSLAFILADAGFDVWNGSRGNTWARRNLYSPDSVEFWAFSF 124
QY 124 DEMARFDPAVINFILOKTGOEKIYVYVGSQGTMGFIAPFTMPPELAOKIKMYFALAPIA 183
DB 125 DEMAKYDLPATIDFILKKTQDKLHVYVHSGQTTIGFIAPFTNPKLAKIKTFYALAPVA 184
QY 184 TVKHAKSPGKTFLLPDMMIKGLFGKKEFLYQTRFLRQ-LVIYLCQGVILDOICSNIMLL 242
DB 185 TVKYTETLLNKLMLVPSFLFKLIFGNKIF-YPHHFFDQFLATEVCSRETVDLLCSNALFI 243
QY 243 LGGFTNNMMSRASVYAAHTLAGTSVQNIILHWSQAVNSGELRAPDWGSETKNLEKCNOP 302
DB 244 ICGFTDNLNMSRLDVLNHPASTSVQNVLHWSQAVKSGKQADPWGSPVQNMHYSQS 303
QY 303 TPVRYVRDMVTPTAMWTGGQDWLSNPEDYKMLLSEVTNLIYHKNIPEWAHVDFIWGLDA 362
DB 304 MPYYNLTDMHVPYIAVWNGGNDLLADPHDVLNLSKLPNLIYHRKIPPNYHLDFIWAMDA 363
QY 363 PHEMYNEIHLM 374
DB 364 PQAVYNEIVSM 375

RESULT 12


```

US-09-186-489-2
; Sequence 2, Application US/09186489
; Patent No. 6375947
; GENERAL INFORMATION:
; APPLICANT: Bolen, Paul L.
; APPLICANT: Cihak, Paul L.
; APPLICANT: Scharpf Jr., Lewis G.
; TITLE OF INVENTION: Purified Recombinant Kid Pregastric Esterase, and
; FILE REFERENCE: Processes for its Production and Use
; CURRENT APPLICATION NUMBER: US/09/186,489
; CURRENT FILING DATE: 1998-11-05
; NUMBER OF SEQ ID NOS: 7
; SOFTWARE: Patent In Ver. 2.0
; SEQ ID NO 2
; LENGTH: 378
; TYPE: PRT
; ORGANISM: Kid (Goat)
; ORGANISM: Kid (Goat)
US-09-186-489-2
Query Match 54.3%; Score 1127; DB 4; Length 378;
Best Local Similarity 55.1%; Pred. No. 1.4e-112;
Matches 205; Conservative 66; Mismatches 101; Indels 0; Gaps 0;

Qy 7 AVDPFAFNISEIHOHQYPCPEEVEVATEDGYILSVNRIPRGLVQPKTGSRRPVLLQHG 66
Db 6 AKNPEASNNVSMISFWGYPSEMKVITADGYILQVYRIPHGKNDANHLQGRPVVFLQHG 65
Qy 67 LVGASNNISNLPNNLSGLFIADAGFDVWMCNSRGNAWSRKHKTLSIDQDEFWAFSDDEM 126
Db 66 LLASATNWSNLPNNLSGLFIADAGFDVWMCNSRGNTWAQEHLYSPDSPEFWAFSDDEM 125
Qy 127 ARFLDPAVINFLQKGTQOEKLYVYVGSQGTMTGFIAPSTMPDLAQKIMYFALAPIATVK 186
Db 126 AEYDLPTSIDFILKRTGQKLYVGHSGQGTIGVFASFSTNPTLAEKIEVHALAPVATVK 185
Qy 187 HAKSPGTFKLLPDMWIKGLFGKEFLYQTRFLRQVLYLQGVQVILDOICSNIMLLGGF 246
Db 186 HTQSLFNKLALIPFLPKIIFGNKMFYHPNFEEQFLGVCSRETDLVLCNKLFAITGA 245
Qy 247 NTNNNMSRASVYAAHTLAGTSVQNLHWSQAVNSGELRAFDWGSSETKNLEKCNQPTPVR 306
Db 246 DNKNFNMSRLDYYVAHNPAGASVQNLHWRQAISGKGFQAFDWGASVENLMHYNQPTPPI 305
Qy 307 YRVDMTPTTAMWTGGQDLNPNEDVKNLSEVTLNLYHKNIPFWAHVDFIWLGLDAPHRM 366
Db 306 YNLTA MVPIAVWSAGQDLADPDQVDLLSKLSNLHKEIPNYNHLDFIWMADAPQEV 365
Qy 367 YNEIHLMQOBE 378
Db 366 YNEIISLMXDK 377

RESULT 13
US-10-043-665B-2
; Sequence 2, Application US/10043665B
; Patent No. 6582948
; GENERAL INFORMATION:
; APPLICANT: Bolen, Paul L.
; APPLICANT: Cihak, Paul L.
; APPLICANT: Scharpf Jr., Lewis G.
; TITLE OF INVENTION: Recombinant Kid Pregastric Esterase and Methods for Its
; FILE REFERENCE: Production and Use
; CURRENT APPLICATION NUMBER: US/10/043,665B
; CURRENT FILING DATE: 2002-09-24
; PRIOR APPLICATION NUMBER: US 09/186,489
; PRIOR FILING DATE: 1998-11-05
; NUMBER OF SEQ ID NOS: 7
; SOFTWARE: Patent In Ver. 2.0
; SEQ ID NO 2
; LENGTH: 378
; TYPE: PRT

```

```

; ORGANISM: Kid (Goat)
US-10-043-665B-2
Query Match 54.3%; Score 1127; DB 4; Length 378;
Best Local Similarity 55.1%; Pred. No. 1.4e-112;
Matches 205; Conservative 66; Mismatches 101; Indels 0; Gaps 0;

Qy 7 AVDPFAFNISEIHOHQYPCPEEVEVATEDGYILSVNRIPRGLVQPKTGSRRPVLLQHG 66
Db 6 AKNPEASNNVSMISFWGYPSEMKVITADGYILQVYRIPHGKNDANHLQGRPVVFLQHG 65
Qy 67 LVGASNNISNLPNNLSGLFIADAGFDVWMCNSRGNAWSRKHKTLSIDQDEFWAFSDDEM 126
Db 66 LLASATNWSNLPNNLSGLFIADAGFDVWMCNSRGNTWAQEHLYSPDSPEFWAFSDDEM 125
Qy 127 ARFLDPAVINFLQKGTQOEKLYVYVGSQGTMTGFIAPSTMPDLAQKIMYFALAPIATVK 186
Db 126 AEYDLPTSIDFILKRTGQKLYVGHSGQGTIGVFASFSTNPTLAEKIEVHALAPVATVK 185
Qy 187 HAKSPGTFKLLPDMWIKGLFGKEFLYQTRFLRQVLYLQGVQVILDOICSNIMLLGGF 246
Db 186 HTQSLFNKLALIPFLPKIIFGNKMFYHPNFEEQFLGVCSRETDLVLCNKLFAITGA 245
Qy 247 NTNNNMSRASVYAAHTLAGTSVQNLHWSQAVNSGELRAFDWGSSETKNLEKCNQPTPVR 306
Db 246 DNKNFNMSRLDYYVAHNPAGASVQNLHWRQAISGKGFQAFDWGASVENLMHYNQPTPPI 305
Qy 307 YRVDMTPTTAMWTGGQDLNPNEDVKNLSEVTLNLYHKNIPFWAHVDFIWLGLDAPHRM 366
Db 306 YNLTA MVPIAVWSAGQDLADPDQVDLLSKLSNLHKEIPNYNHLDFIWMADAPQEV 365
Qy 367 YNEIHLMQOBE 378
Db 366 YNEIISLMXDK 377

RESULT 14
US-09-348-930A-9
; Sequence 9, Application US/09348930A
; Patent No. 6573431
; GENERAL INFORMATION:
; APPLICANT: Lenee, P
; APPLICANT: Gruber, V
; APPLICANT: Baudino, S
; APPLICANT: Merot, B
; APPLICANT: Benicourt, C
; APPLICANT: Cudrey, C
; TITLE OF INVENTION: Recombinant Preduodenal Lipases and Polypeptides Derivatives Prod
; FILE REFERENCE: Plants, Processes for Obtaining Them and Their Uses
; CURRENT APPLICATION NUMBER: US/09/348,930A
; CURRENT FILING DATE: 1999-07-02
; PRIOR APPLICATION NUMBER: 08/945,321
; PRIOR FILING DATE: 1998-02-12
; PRIOR APPLICATION NUMBER: PCT/FR96/00606
; PRIOR FILING DATE: 1996-03-19
; NUMBER OF SEQ ID NOS: 37
; SOFTWARE: Patent in version 3.0
; SEQ ID NO 9
; LENGTH: 398
; TYPE: PRT
; ORGANISM: Homo sapiens
; ORGANISM: Homo sapiens
US-09-348-930A-9
Query Match 53.4%; Score 1109; DB 4; Length 398;
Best Local Similarity 54.7%; Pred. No. 1.3e-110;
Matches 202; Conservative 69; Mismatches 98; Indels 0; Gaps 0;

Qy 10 PEAFNISEIHOHQYPCPEEVEVATEDGYILSVNRIPRGLVQPKTGSRRPVLLQHG 69
Db 29 PEVTNISMITYGYPNEEVEVATEDGYILEVNRIPYKKNKSGNTGQRPVVFVFLQGLLA 88
Qy 70 GASNNISNLPNNLSGLFIADAGFDVWMCNSRGNAWSRKHKTLSIDQDEFWAFSDDEM 129

```

```

Db      89 SATNWSLNNPNSLAFILADAGYVWLGNSRGNTWARRNLYYSPDSVEFWAFSDEWAKY 148
QY      130 DLPVAVINFLQKTGQEKIYYVGYSGTWTMGFTAFSTMPBLAQKIYKYPALAPIATVGHAK 189
Db      149 DLPATIDFIVKTKGQKQLHYVGHSGTWTMGFTAFSTNPSLAKRIKTFYALAPVATVKYTK 208
QY      190 SPGTFLLLPDMNMIKGLFGKKEFLYQTRFLROLVYLCGOVILDOICSNIMLLGGFWTN 249
Db      209 SLINKLRFVPSQLFKFIDGKLFYPHNFDQFLATEVCSREMLNLLCSNALFIICGFSK 268
QY      250 NNMGRASYAARTLAGTSVQNILHWSQAVNSGELRAFDMGSETKNLEKCKNQPTPVRYR 309
Db      269 NFNTSLDVLVLSHNPAGTSVQNMFWHTQAVKSGKQAYDWGSPVQNRMHYDQSQPPYYNV 328
QY      310 RDMVTPTAMWTGGQWLSNPEDVKMLSEVNTLIYHKNIPEWAHVDFIWLGLDAPHERMNE 369
Db      329 TANNVPIAVWNGGKOLLADPDQVGLLLPKLPNLIIYHKEIPFYNHLLDFIWMADAPQEVND 388
QY      370 IHLMOQEE 378
Db      389 IVSMISEDK 397

```

RESULT 15

```

US-08-227-108-16
; Sequence 16, Application US/08227108
; Patent No. 5807726
; GENERAL INFORMATION:
; APPLICANT: Blanchard, Claire
; APPLICANT: Benicourt, Claude
; APPLICANT: Junien, Jean-Louis
; TITLE OF INVENTION: Recombinant Dog Gastric Lipase
; NUMBER OF SEQUENCES: 21
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Pennie & Edmonds
; STREET: 1155 Avenue of the Americas
; CITY: New York
; STATE: New York
; COUNTRY: U.S.A.
; ZIP: 10036
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/227,108
; FILING DATE: 03-APR-1994
; CLASSIFICATION: 435
; ATTORNEY/AGENT INFORMATION:
; NAME: Panucci, Allan A.
; REGISTRATION NUMBER: 30,256
; REFERENCE/DOCKET NUMBER: 7620-033
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 212 790-9090
; TELEFAX: 212 869-8864/9741
; TELEX: 66141 PENNIE
; INFORMATION FOR SEQ ID NO: 16:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 380 amino acids
; TYPE: amino acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: protein
US-08-227-108-16

```

```

Query Match      52.9%; Score 1098.5; DB 1; Length 380;
Best Local Similarity 54.6%; Pred. No. 1.6e-109;
Matches 202; Conservative 69; Mismatches 98; Indels 1; Gaps 1;
QY      10 PEAFMNISEIIQHGYCEYEYVATEDGYILSVNRIPIRGVQPKTKGSRPVLLQHLGLVG 69
      || |||||:|: ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||

```

```

Db      10 PEVTMNSQMITWGYPNEEYEVWTDGYILEVNRIPYVGKNSGNTGQRPVVFLQHLGLA 69
QY      70 GASNWSLNNPNSLGFILADAGYVWLGNSRGNTWARRNLYYSPDSVEFWAFSDEWAK 128
Db      70 SATNWSLNNPNSLAFILADAGYVWLGNSRGNTWARRNLYYSPDSVEFWAFSDEWAK 129
QY      129 FDLPAVINFLQKTGQEKIYYVGYSGTWTMGFTAFSTMPBLAQKIYKYPALAPIATVGHAK 188
Db      130 YDLPATIDFIVKTKGQKQLHYVGHSGTWTMGFTAFSTNPSLAKRIKTFYALAPVATVKY 189
QY      189 KSPGTFLLLPDMNMIKGLFGKKEFLYQTRFLROLVYLCGOVILDOICSNIMLLGGFWTN 248
Db      190 KSLINKLRFVPSQLFKFIDGKLFYPHNFDQFLATEVCSREMLNLLCSNALFIICGFSK 249
QY      249 NNMGRASYAARTLAGTSVQNILHWSQAVNSGELRAFDMGSETKNLEKCKNQPTPVRYR 308
Db      250 NFNTSLDVLVLSHNPAGTSVQNMFWHTQAVKSGKQAYDWGSPVQNRMHYDQSQPPYYNV 309
QY      309 RDMVTPTAMWTGGQWLSNPEDVKMLSEVNTLIYHKNIPEWAHVDFIWLGLDAPHERMNE 368
Db      310 TANNVPIAVWNGGKOLLADPDQVGLLLPKLPNLIIYHKEIPFYNHLLDFIWMADAPQEVND 369
QY      369 IHLMOQEE 378
Db      370 IVSMISEDK 379

```

Search completed: February 19, 2004, 14:23:48
Job time : 15.4317 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2004 Compugen Ltd.

OM protein - protein search, using sw model

Run on: February 19, 2004, 14:22:29 ; Search time 28.6162 Seconds
(without alignments)
3095.051 Million cell updates/sec

Title: US-10-042-431-47
Perfect score: 2247
Sequence: 1 MLETLRSQIVSHRMEMWLL.....IHLMQEETNLQGRCEAVL 423

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 801455 seqs, 209382283 residues

Total number of hits satisfying chosen parameters: 801455

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : Published Applications AA:
1: /cgn2_6/ptodata/1/pubpaa/US07_PUBCOMB.pep.*
2: /cgn2_6/ptodata/1/pubpaa/PCT_NEW_PUB.pep.*
3: /cgn2_6/ptodata/1/pubpaa/US06_NEW_PUB.pep.*
4: /cgn2_6/ptodata/1/pubpaa/US06_PUBCOMB.pep.*
5: /cgn2_6/ptodata/1/pubpaa/US07_NEW_PUB.pep.*
6: /cgn2_6/ptodata/1/pubpaa/PCTUS_PUBCOMB.pep.*
7: /cgn2_6/ptodata/1/pubpaa/US08_NEW_PUB.pep.*
8: /cgn2_6/ptodata/1/pubpaa/US08_PUBCOMB.pep.*
9: /cgn2_6/ptodata/1/pubpaa/US09A_PUBCOMB.pep.*
10: /cgn2_6/ptodata/1/pubpaa/US09B_PUBCOMB.pep.*
11: /cgn2_6/ptodata/1/pubpaa/US09C_PUBCOMB.pep.*
12: /cgn2_6/ptodata/1/pubpaa/US09_NEW_PUB.pep.*
13: /cgn2_6/ptodata/1/pubpaa/US10A_PUBCOMB.pep.*
14: /cgn2_6/ptodata/1/pubpaa/US10B_PUBCOMB.pep.*
15: /cgn2_6/ptodata/1/pubpaa/US10C_PUBCOMB.pep.*
16: /cgn2_6/ptodata/1/pubpaa/US10_NEW_PUB.pep.*
17: /cgn2_6/ptodata/1/pubpaa/US60_NEW_PUB.pep.*
18: /cgn2_6/ptodata/1/pubpaa/US60_PUBCOMB.pep.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	2247	100.0	423	11	US-09-759-1308-417
2	2247	100.0	423	14	US-10-042-431-47
3	2174	96.8	409	10	US-09-835-996A-21
4	2076	92.4	390	11	US-09-759-1308-419
5	2076	92.4	390	14	US-10-042-431-49
6	1991	88.6	374	10	US-09-811-825-4
7	1991	88.6	398	10	US-09-811-825-2
8	1290	57.4	399	12	US-10-051-874-93
9	1289	57.4	399	12	US-10-051-874-94
10	1289	57.4	399	11	US-09-759-1308-411
11	1289	57.4	399	12	US-10-051-874-92
12	1289	57.4	399	14	US-10-042-431-41
13	1282	57.1	392	14	US-10-003-302-4
14	1165	51.8	221	11	US-09-759-1308-420
15	1165	51.8	221	14	US-10-042-431-50

16	1164.5	51.8	395	12	US-10-015-115-98
17	1154	51.4	398	12	US-10-015-115-96
18	1141	50.8	357	12	US-10-108-260A-3705
19	1139	50.7	399	12	US-10-221-097-35
20	1136	50.6	397	12	US-10-312-088-29
21	1127	50.2	378	12	US-10-420-564-2
22	1119	49.8	398	11	US-09-759-130B-445
23	1119	49.8	398	12	US-10-015-115-94
24	1119	49.8	398	12	US-10-015-115-95
25	1119	49.8	398	14	US-10-042-431-75
26	1109	49.4	371	12	US-10-015-115-97
27	1090	48.5	398	14	US-10-056-744B-2
28	1034	46.0	401	12	US-10-015-115-32
29	1031	45.9	395	14	US-10-003-302-2
30	1006	44.8	373	12	US-10-051-874-28
31	1003	44.6	373	12	US-10-051-874-30
32	977	43.5	365	12	US-10-312-088-28
33	977	43.5	395	12	US-10-051-874-90
34	937	41.7	349	12	US-10-051-874-32
35	901.5	40.1	390	12	US-10-015-115-30
36	843	37.5	351	12	US-10-051-874-91
37	783	34.8	144	11	US-09-759-130B-422
38	783	34.8	144	14	US-10-042-431-52
39	400	17.8	127	14	US-10-056-744B-5
40	319	14.2	567	12	US-10-032-585-7720
41	301	13.4	476	15	US-10-128-714-3075
42	301	13.4	641	15	US-10-128-714-8075
43	294	13.1	431	12	US-10-304-038-6
44	209	9.3	86	12	US-09-864-408A-8008
45	171	7.6	33	11	US-09-759-130B-418

ALIGNMENTS

RESULT 1

US-09-759-1308-417
; Sequence 417, Application US/09759130B
; Publication No. US20030022279A1
; GENERAL INFORMATION:
; APPLICANT: Millennium Pharmaceuticals, Inc.
; APPLICANT: McCarthy, Sean A
; APPLICANT: Fraser, Christopher C
; APPLICANT: Sharp, John D
; APPLICANT: Barnes, Thomas S
; APPLICANT: Kirst, Susan J
; APPLICANT: Mackay, Charles R
; APPLICANT: Myers, Paul S
; APPLICANT: Leiby, Kevin R
; APPLICANT: Wrighton, Nicolas
; APPLICANT: Goodearl, Andrew
; APPLICANT: Holtzman, Douglas A
; TITLE OF INVENTION: NOVEL GENES ENCODING PROTEINS HAVING
; TITLE OF INVENTION: PROGNOSTIC, DIAGNOSTIC, PREVENTIVE, THERAPEUTIC, AND OTHER
; TITLE OF INVENTION: USES.
; FILE REFERENCE: MP100-5350MIM
; CURRENT APPLICATION NUMBER: US/09/759,130B
; CURRENT FILING DATE: 2002-09-16
; PRIOR APPLICATION NUMBER: US 09/479,249
; PRIOR FILING DATE: 2000-01-07
; PRIOR APPLICATION NUMBER: US 09/559,497
; PRIOR FILING DATE: 2000-04-27
; PRIOR APPLICATION NUMBER: US 09/578,063
; PRIOR FILING DATE: 2000-05-24
; PRIOR APPLICATION NUMBER: US 09/333,159
; PRIOR FILING DATE: 1999-06-14
; PRIOR APPLICATION NUMBER: US 09/596,194
; PRIOR FILING DATE: 2000-07-14
; PRIOR APPLICATION NUMBER: US 09/342,364
; PRIOR FILING DATE: 1999-06-29
; PRIOR APPLICATION NUMBER: US 09/608,452
; PRIOR FILING DATE: 2000-06-30
; PRIOR APPLICATION NUMBER: US 09/393,996

```

; PRIOR FILING DATE: 1999-09-10
; PRIOR APPLICATION NUMBER: US 09/602,871
; PRIOR FILING DATE: 2000-06-23
; PRIOR APPLICATION NUMBER: US 09/420,707
; PRIOR FILING DATE: 1999-10-19
; NUMBER OF SEQ ID NOS: 460
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 417
; LENGTH: 423
; TYPE: PRT
; ORGANISM: Homo sapiens
; US-09-759-1309-417

Query Match      100.0%; Score 2247; DB 11; Length 423;
Best Local Similarity 100.0%; Pred. No. 2.6e-229;
Matches 423; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 MLETLRQWIVSHRMEMWLLILVAYMFQNVNSVHMPKAVDPEAFMNISEIIHQGYPC 60
DB 1 MLETLRQWIVSHRMEMWLLILVAYMFQNVNSVHMPKAVDPEAFMNISEIIHQGYPC 60
QY 61 EYEVEATEDGYILSVNRI PRGLVOPKKTGSRPVVLLQHGLVGGASNWI SNL PNNLSLGFIL 120
DB 61 EYEVEATEDGYILSVNRI PRGLVOPKKTGSRPVVLLQHGLVGGASNWI SNL PNNLSLGFIL 120
QY 121 ADAGFDVWNGSNGNWSRKHKTLSIDQDEFWAFSYDENARFDLPVAVINFILOKGTQGEKI 180
DB 121 ADAGFDVWNGSNGNWSRKHKTLSIDQDEFWAFSYDENARFDLPVAVINFILOKGTQGEKI 180
QY 181 YYVGYSGQTTMGFIATSTWPELAQKIKMYFALAPIATVVKHAKSPGTFKLLPDMMIKGLF 240
DB 181 YYVGYSGQTTMGFIATSTWPELAQKIKMYFALAPIATVVKHAKSPGTFKLLPDMMIKGLF 240
QY 241 GKKEFLYQTRFLRQLVYLCGVILDOICSNIMLLGGFTNNMNSRASVYAAHTLAGT 300
DB 241 GKKEFLYQTRFLRQLVYLCGVILDOICSNIMLLGGFTNNMNSRASVYAAHTLAGT 300
QY 301 SVQNILHWSQAVNSGELRAPDWSGSETKNLEKCNQPTPVRYRVRDMDTPTAMWTGGQDWLS 360
DB 301 SVQNILHWSQAVNSGELRAPDWSGSETKNLEKCNQPTPVRYRVRDMDTPTAMWTGGQDWLS 360
QY 361 NPEDVKMLSEVNTLIYHKNIPEWAHVDFTWGLDAPHRMVNEIHLMOQEETNLSQGRCE 420
DB 361 NPEDVKMLSEVNTLIYHKNIPEWAHVDFTWGLDAPHRMVNEIHLMOQEETNLSQGRCE 420
QY 421 AVL 423
DB 421 AVL 423

```

```

RESULT 2
US-10-042-431-47
; Sequence 47, Application US/10042431
; Publication No. US20020182675A1
; GENERAL INFORMATION:
; APPLICANT: MCCARTHY, Sean A
; APPLICANT: BARNES, Thomas M
; APPLICANT: FRASER, Christopher C
; APPLICANT: SHARP, John D
; TITLE OF INVENTION: NOVEL GENES ENCODING PROTEINS HAVING DIAGNOSTIC,
; PREVENTIVE, THERAPEUTIC, AND OTHER USES
; FILE REFERENCE: 10147-6U2
; CURRENT FILING DATE: 2001-10-25
; PRIOR APPLICATION NUMBER: US 09/333,159
; PRIOR FILING DATE: 1999-06-14
; PRIOR APPLICATION NUMBER: US 09/578,063
; PRIOR FILING DATE: 2000-05-24
; NUMBER OF SEQ ID NOS: 79
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 47
; LENGTH: 423
; TYPE: PRT

```

```

; ORGANISM: Homo sapiens
; US-10-042-431-47

Query Match      100.0%; Score 2247; DB 14; Length 423;
Best Local Similarity 100.0%; Pred. No. 2.6e-229;
Matches 423; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 MLETLRQWIVSHRMEMWLLILVAYMFQNVNSVHMPKAVDPEAFMNISEIIHQGYPC 60
DB 1 MLETLRQWIVSHRMEMWLLILVAYMFQNVNSVHMPKAVDPEAFMNISEIIHQGYPC 60
QY 61 EYEVEATEDGYILSVNRI PRGLVOPKKTGSRPVVLLQHGLVGGASNWI SNL PNNLSLGFIL 120
DB 61 EYEVEATEDGYILSVNRI PRGLVOPKKTGSRPVVLLQHGLVGGASNWI SNL PNNLSLGFIL 120
QY 121 ADAGFDVWNGSNGNWSRKHKTLSIDQDEFWAFSYDENARFDLPVAVINFILOKGTQGEKI 180
DB 121 ADAGFDVWNGSNGNWSRKHKTLSIDQDEFWAFSYDENARFDLPVAVINFILOKGTQGEKI 180
QY 181 YYVGYSGQTTMGFIATSTWPELAQKIKMYFALAPIATVVKHAKSPGTFKLLPDMMIKGLF 240
DB 181 YYVGYSGQTTMGFIATSTWPELAQKIKMYFALAPIATVVKHAKSPGTFKLLPDMMIKGLF 240
QY 241 GKKEFLYQTRFLRQLVYLCGVILDOICSNIMLLGGFTNNMNSRASVYAAHTLAGT 300
DB 241 GKKEFLYQTRFLRQLVYLCGVILDOICSNIMLLGGFTNNMNSRASVYAAHTLAGT 300
QY 301 SVQNILHWSQAVNSGELRAPDWSGSETKNLEKCNQPTPVRYRVRDMDTPTAMWTGGQDWLS 360
DB 301 SVQNILHWSQAVNSGELRAPDWSGSETKNLEKCNQPTPVRYRVRDMDTPTAMWTGGQDWLS 360
QY 361 NPEDVKMLSEVNTLIYHKNIPEWAHVDFTWGLDAPHRMVNEIHLMOQEETNLSQGRCE 420
DB 361 NPEDVKMLSEVNTLIYHKNIPEWAHVDFTWGLDAPHRMVNEIHLMOQEETNLSQGRCE 420
QY 421 AVL 423
DB 421 AVL 423

RESULT 3
US-09-835-996A-21
; Sequence 21, Application US/09835996A
; Patent No. US20020142953A1
; GENERAL INFORMATION:
; APPLICANT: Ballinger, Dennis
; APPLICANT: Loeb, Debra
; APPLICANT: Montgomery, Julie
; APPLICANT: Tang, Y. Tom
; APPLICANT: Zhou, Ping
; APPLICANT: Goodrich, Ryle
; APPLICANT: Liu, Chenghua
; APPLICANT: Asundi, Vinod
; APPLICANT: Zhao, Qing
; APPLICANT: Wehrman, Tom
; APPLICANT: Drmanac, Radoje
; APPLICANT: Ren, Feiyan
; APPLICANT: Qian, Xishong
; APPLICANT: Wang, Durrui
; TITLE OF INVENTION: MATERIALS AND METHODS RELATING TO LIPID METABOLISM
; FILE REFERENCE: 28110/35915A
; CURRENT APPLICATION NUMBER: US/09/835,996A
; CURRENT FILING DATE: 2001-04-16
; PRIOR APPLICATION NUMBER: US 60/197,137
; PRIOR FILING DATE: 2000-04-14
; PRIOR APPLICATION NUMBER: US 09/714,936
; PRIOR FILING DATE: 2000-11-17
; PRIOR APPLICATION NUMBER: US 09/667,298
; PRIOR FILING DATE: 2000-09-22
; PRIOR APPLICATION NUMBER: US 09/631,451
; PRIOR FILING DATE: 2000-08-03
; PRIOR APPLICATION NUMBER: US 09/598,042
; PRIOR FILING DATE: 2000-06-20

```

```
; NUMBER OF SEQ ID NOS: 45
; SOFTWARE: PatentIn version 3.0
; SEQ ID NO 21
; LENGTH: 409
; TYPE: PRT
; ORGANISM: Homo sapiens
US-09-835-996A-21

Query Match      96.8%; Score 2174; DB 10; Length 409;
Best Local Similarity 100.0%; Pred. No. 1.4e-221; Indels 0; Gaps 0;
Matches 409; Conservative 0; Mismatches 0;

QY 15 MEMLLILVYMFQNVNSVHMPKAVDPPEAFMNISEIIHQGYPCBEYEYVATEDGYLS 74
DB 1 MEMLLILVYMFQNVNSVHMPKAVDPPEAFMNISEIIHQGYPCBEYEYVATEDGYLS 60
QY 75 VNIIPRGVLVQPKTGSRRPVLLQHLGVGASNWIENLNNLSGLFILADAGFDVWNGSRG 134
DB 61 VNIIPRGVLVQPKTGSRRPVLLQHLGVGASNWIENLNNLSGLFILADAGFDVWNGSRG 120
QY 135 NAWSRKHKTLSIDQDEFWAFSYDEMAREFDLPVAVINFILOKTQGEKIYYVGYSGQTTMGFI 194
DB 121 NAWSRKHKTLSIDQDEFWAFSYDEMAREFDLPVAVINFILOKTQGEKIYYVGYSGQTTMGFI 180
QY 195 AFSTWPELAQKIKMYFALAPATVGHAKSPGTFKLLIPDMMIKGLFGKKEFLYQTRFLRQ 254
DB 181 AFSTWPELAQKIKMYFALAPATVGHAKSPGTFKLLIPDMMIKGLFGKKEFLYQTRFLRQ 240
QY 255 LVYLICGQVILDQICSNIMLLIGGFNTNNMNSRASVVAHTLAGTSVQNTILHWSQAVNS 314
DB 241 LVYLICGQVILDQICSNIMLLIGGFNTNNMNSRASVVAHTLAGTSVQNTILHWSQAVNS 300
QY 315 GELRAFDSGETKNEKCNQPTPVRYRVRDVTPTAMTGGQDMLSNPDEVKMLLSEVTN 374
DB 301 GELRAFDSGETKNEKCNQPTPVRYRVRDVTPTAMTGGQDMLSNPDEVKMLLSEVTN 360
QY 375 LIYHKNIPENAHVDFIWLGLDAPHRMYNEIILHMQOEEETNLSQGRCEAVL 423
DB 361 LIYHKNIPENAHVDFIWLGLDAPHRMYNEIILHMQOEEETNLSQGRCEAVL 409

RESULT 4
US-09-759-130B-419
; Sequence 419, Application US/09759130B
; Publication No. US2003002279A1
; GENERAL INFORMATION:
; APPLICANT: Millennium Pharmaceuticals, Inc.
; APPLICANT: McCarthy, Sean A
; APPLICANT: Fraser, Christopher C
; APPLICANT: Sharp, John D
; APPLICANT: Barnes, Thomas S
; APPLICANT: Kirt, Susan J
; APPLICANT: Mackay, Charles R
; APPLICANT: Myers, Paul S
; APPLICANT: Leiby, Kevin R
; APPLICANT: Wrighton, Nicolas
; APPLICANT: Goodearl, Andrew
; APPLICANT: Holtzman, Douglas A
; TITLE OF INVENTION: NOVEL GENES ENCODING PROTEINS HAVING
; TITLE OF INVENTION: PROGNOSTIC, DIAGNOSTIC, PREVENTIVE, THERAPEUTIC, AND OTHER
; FILE REFERENCE: MP100-5350NIM
; CURRENT APPLICATION NUMBER: US/09/759,130B
; CURRENT FILING DATE: 2002-09-16
; PRIOR APPLICATION NUMBER: US/09/479,249
; PRIOR FILING DATE: 2000-01-07
; PRIOR APPLICATION NUMBER: US/09/559,497
; PRIOR FILING DATE: 2000-04-27
; PRIOR APPLICATION NUMBER: US/09/578,063
; PRIOR FILING DATE: 2000-05-24
; PRIOR APPLICATION NUMBER: US/09/333,159
; PRIOR FILING DATE: 1999-06-14
; PRIOR APPLICATION NUMBER: US/09/596,194
```

```
; PRIOR FILING DATE: 2000-07-14
; PRIOR APPLICATION NUMBER: US/09/342,364
; PRIOR FILING DATE: 1999-06-29
; PRIOR APPLICATION NUMBER: US/09/608,452
; PRIOR FILING DATE: 2000-06-30
; PRIOR APPLICATION NUMBER: US/09/393,996
; PRIOR FILING DATE: 1999-09-10
; PRIOR APPLICATION NUMBER: US/09/602,871
; PRIOR FILING DATE: 2000-06-23
; PRIOR APPLICATION NUMBER: US/09/420,707
; PRIOR FILING DATE: 1999-10-19
; NUMBER OF SEQ ID NOS: 460
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 419
; LENGTH: 390
; TYPE: PRT
; ORGANISM: Homo sapiens
US-09-759-130B-419

Query Match      92.4%; Score 2076; DB 11; Length 390;
Best Local Similarity 100.0%; Pred. No. 3.2e-211; Indels 0; Gaps 0;
Matches 390; Conservative 0; Mismatches 0;

QY 34 VHMPTKAVDPPEAFMNISEIIHQGYPCBEYEYVATEDGYLSVNIIPRGVLVQPKTGSRRV 93
DB 1 VHMPTKAVDPPEAFMNISEIIHQGYPCBEYEYVATEDGYLSVNIIPRGVLVQPKTGSRRV 60
QY 94 VLIQHLGVGASNWIENLNNLSGLFILADAGFDVWNGSRGNWRSRKHKTLSIDQDEFWA 153
DB 61 VLIQHLGVGASNWIENLNNLSGLFILADAGFDVWNGSRGNWRSRKHKTLSIDQDEFWA 120
QY 154 FSYDEMAREFDLPVAVINFILOKTQGEKIYYVGYSGQTTMGFIASFSTWPELAQKIKMYFALA 213
DB 121 FSYDEMAREFDLPVAVINFILOKTQGEKIYYVGYSGQTTMGFIASFSTWPELAQKIKMYFALA 180
QY 214 PIATVGHAKSPGTFKLLIPDMMIKGLFGKKEFLYQTRFLRQLVYLICGQVILDQICSNIM 273
DB 181 PIATVGHAKSPGTFKLLIPDMMIKGLFGKKEFLYQTRFLRQLVYLICGQVILDQICSNIM 240
QY 274 LLLGGFNTNNMNSRASVVAHTLAGTSVQNTILHWSQAVNSGELRAFDSGETKNEKCN 333
DB 241 LLLGGFNTNNMNSRASVVAHTLAGTSVQNTILHWSQAVNSGELRAFDSGETKNEKCN 300
QY 334 QPTPVRYRVRDVTPTAMTGGQDMLSNPDEVKMLLSEVTNLIYHKNIPENAHVDFIWLGL 393
DB 301 QPTPVRYRVRDVTPTAMTGGQDMLSNPDEVKMLLSEVTNLIYHKNIPENAHVDFIWLGL 360
QY 394 DAPHRMYNEIILHMQOEEETNLSQGRCEAVL 423
DB 361 DAPHRMYNEIILHMQOEEETNLSQGRCEAVL 390

RESULT 5
US-10-042-431-49
; Sequence 49, Application US/10042431
; Publication No. US20020182675A1
; GENERAL INFORMATION:
; APPLICANT: MCCARTHY, Sean A
; APPLICANT: BARNES, Thomas M
; APPLICANT: FRASER, Christopher C
; APPLICANT: SHARP, John D
; TITLE OF INVENTION: NOVEL GENES ENCODING PROTEINS HAVING DIAGNOSTIC,
; TITLE OF INVENTION: PREVENTIVE, THERAPEUTIC, AND OTHER USES
; FILE REFERENCE: 10147-6U2
; CURRENT APPLICATION NUMBER: US/10/042,431
; CURRENT FILING DATE: 2001-10-25
; PRIOR APPLICATION NUMBER: US/09/333,159
; PRIOR FILING DATE: 1999-06-14
; PRIOR APPLICATION NUMBER: US/09/578,063
; PRIOR FILING DATE: 2000-05-24
; NUMBER OF SEQ ID NOS: 79
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 49
```

```

; LENGTH: 390
; TYPE: PRT
; ORGANISM: Homo sapiens
US-10-042-431-49

Query Match      92.4%; Score 2076; DB 14; Length 390;
Best Local Similarity 100.0%; Pred. No. 3.2e-211;
Matches 390; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 34 VHMPTKAVDPEAFMNISEIIHQGYPCSEYEYVATEDGYILSVNRIPRGLVQPKTGSRPV 93
DB 1 VHMPTKAVDPEAFMNISEIIHQGYPCSEYEYVATEDGYILSVNRIPRGLVQPKTGSRPV 60

QY 94 VLLQHLGLVGASNIWISNLPNNSLGFILADAGFDVWNGSRGNWAKSKHKTLSIDQDEFWA 153
DB 61 VLLQHLGLVGASNIWISNLPNNSLGFILADAGFDVWNGSRGNWAKSKHKTLSIDQDEFWA 120

QY 154 PSYDEMARFDPRAVINFILOKTQOEKIYYGVYSGTGTMGFIAPSTMPPELAQIKMYFALA 213
DB 121 PSYDEMARFDPRAVINFILOKTQOEKIYYGVYSGTGTMGFIAPSTMPPELAQIKMYFALA 180

QY 214 PIATVXHAKEGPGTKFLLPDMMIKGLFGKKEFLYQTRFLRQLVYLCOGVILDOICSNIM 273
DB 181 PIATVXHAKEGPGTKFLLPDMMIKGLFGKKEFLYQTRFLRQLVYLCOGVILDOICSNIM 240

QY 274 LLLGGFTNNMNSRASVVAHTLAGTSVQNILHWSQAVNSGELRAFDWGSSEKCNLEKCN 333
DB 241 LLLGGFTNNMNSRASVVAHTLAGTSVQNILHWSQAVNSGELRAFDWGSSEKCNLEKCN 300

QY 334 QPTPVRYRVRDMTVPMTAMWTGGQDMLSNPDEVKMLSEVTNLIYHKNIPWAHVDPIWGL 393
DB 301 QPTPVRYRVRDMTVPMTAMWTGGQDMLSNPDEVKMLSEVTNLIYHKNIPWAHVDPIWGL 360

QY 394 DAPHRMYNEIHLMOQETNLSQRCCEAVL 423
DB 361 DAPHRMYNEIHLMOQETNLSQRCCEAVL 390

```

```

RESULT 6
US-09-811-825-4
; Sequence 4, Application US/09811825
; Patent No. US20020144297A1
; GENERAL INFORMATION:
; APPLICANT: YAN, Chunhua et al.
; TITLE OF INVENTION: ISOLATED HUMAN LIPASE PROTEINS, NUCLEIC
; TITLE OF INVENTION: ACID MOLECULES ENCODING HUMAN LIPASE PROTEINS, AND USES
; TITLE OF INVENTION: THEREOF
; FILE REFERENCE: CL001170
; CURRENT APPLICATION NUMBER: US/09/811,825
; NUMBER OF SEQ ID NOS: 4
; SOFTWARE: Fast-Seq for Windows Version 4.0
; SEQ ID NO 4
; LENGTH: 374
; TYPE: PRT
; ORGANISM: Human
US-09-811-825-4

Query Match      88.6%; Score 1991; DB 10; Length 374;
Best Local Similarity 100.0%; Pred. No. 3.1e-202;
Matches 374; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 50 SEIIHQGYPCSEYEYVATEDGYILSVNRIPRGLVQPKTGSRPVLLQHLVGASNIWIS 109
DB 1 SEIIHQGYPCSEYEYVATEDGYILSVNRIPRGLVQPKTGSRPVLLQHLVGASNIWIS 60

QY 110 NLPNNSLGFILADAGFDVWNGSRGNWAKSKHKTLSIDQDEFWAFSYDEMARFDPRAVIN 169
DB 61 NLPNNSLGFILADAGFDVWNGSRGNWAKSKHKTLSIDQDEFWAFSYDEMARFDPRAVIN 120

QY 170 FILOKTQOEKIYYGVYSGTGTMGFIAPSTMPPELAQIKMYFALAPIATVXHAKEGPGTKF 229
DB 121 FILOKTQOEKIYYGVYSGTGTMGFIAPSTMPPELAQIKMYFALAPIATVXHAKEGPGTKF 180

RESULT 8
US-10-051-874-93

Query Match      88.6%; Score 1991; DB 10; Length 398;
Best Local Similarity 100.0%; Pred. No. 3.4e-202;
Matches 374; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 50 SEIIHQGYPCSEYEYVATEDGYILSVNRIPRGLVQPKTGSRPVLLQHLVGASNIWIS 109
DB 25 SEIIHQGYPCSEYEYVATEDGYILSVNRIPRGLVQPKTGSRPVLLQHLVGASNIWIS 84

QY 110 NLPNNSLGFILADAGFDVWNGSRGNWAKSKHKTLSIDQDEFWAFSYDEMARFDPRAVIN 169
DB 85 NLPNNSLGFILADAGFDVWNGSRGNWAKSKHKTLSIDQDEFWAFSYDEMARFDPRAVIN 144

QY 170 FILOKTQOEKIYYGVYSGTGTMGFIAPSTMPPELAQIKMYFALAPIATVXHAKEGPGTKF 229
DB 145 FILOKTQOEKIYYGVYSGTGTMGFIAPSTMPPELAQIKMYFALAPIATVXHAKEGPGTKF 204

QY 230 LLLPDMMIKGLFGKKEFLYQTRFLRQLVYLCOGVILDOICSNIMLLGGFTNNMNSRAS 289
DB 205 LLLPDMMIKGLFGKKEFLYQTRFLRQLVYLCOGVILDOICSNIMLLGGFTNNMNSRAS 264

QY 290 SVYAAHTLAGTSVQNILHWSQAVNSGELRAFDWGSSEKCNLEKCNQPTPVRYRVRDMTVP 349
DB 265 SVYAAHTLAGTSVQNILHWSQAVNSGELRAFDWGSSEKCNLEKCNQPTPVRYRVRDMTVP 324

QY 350 AMWTGGQDMLSNPDEVKMLSEVTNLIYHKNIPWAHVDPIWGLDAPHRMYNEIHLMOQ 409
DB 325 AMWTGGQDMLSNPDEVKMLSEVTNLIYHKNIPWAHVDPIWGLDAPHRMYNEIHLMOQ 384

QY 410 EETNLSQRCCEAVL 423
DB 385 EETNLSQRCCEAVL 398

```



```

; Sequence 93, Application US/10051874
; Publication No. US20040005557A1
; GENERAL INFORMATION:
; APPLICANT: Padigar, Muralidhara
; APPLICANT: Alsobrook II, John P
; APPLICANT: Colman, Steven D
; APPLICANT: Spytek, Kimberly A
; APPLICANT: Boldog, Ferenc
; APPLICANT: Vernet, Corine AM
; APPLICANT: Li, Li
; APPLICANT: Shenoy, Suresh G
; APPLICANT: Casman, Stacie J
; APPLICANT: Guo, Xiaojia Sasha
; APPLICANT: Edinger, Shlomit R
; APPLICANT: MacDougall, John R
; APPLICANT: Malyankar, Uriel M
; APPLICANT: Patturajan, Meera
; APPLICANT: Shinkets, Richard A
; APPLICANT: Pena, Carol EA
; APPLICANT: Tchernev, Velizar T
; APPLICANT: Zerhusen, Bryan D
; APPLICANT: Millet, Isabelle
; APPLICANT: Miller, Charles E
; APPLICANT: Lepley, Denise M
; APPLICANT: Smithson, Glenda
; APPLICANT: Baumgartner, Jason C
; APPLICANT: Herrman, John L
; APPLICANT: Peyman, John A
; APPLICANT: Gorman, Linda
; APPLICANT: Mezes, Peter D
; APPLICANT: Kekuda, Ramesh
; APPLICANT: Taupier Jr, Raymond J
; APPLICANT: Gerlach, Valerie
; APPLICANT: Grosse, William M
; APPLICANT: Liu, Xiaohong
; APPLICANT: Ellerman, Karen
; APPLICANT: Rothenberg, Mark
; APPLICANT: Stone, David J
; APPLICANT: Burgess, Catherine E
; TITLE OF INVENTION: PROTEINS, POLYNUCLEOTIDES ENCODING THEM AND METHODS OF
; TITLE OF INVENTION: USING THE SAME
; FILE REFERENCE: 21402-245
; CURRENT APPLICATION NUMBER: US/10/051,874
; CURRENT FILING DATE: 2002-09-25
; PRIOR APPLICATION NUMBER: 60/268,595
; PRIOR FILING DATE: 2001-02-14
; PRIOR APPLICATION NUMBER: 60/325,306
; PRIOR FILING DATE: 2001-03-27
; PRIOR APPLICATION NUMBER: 60/262,587
; PRIOR FILING DATE: 2001-01-18
; PRIOR APPLICATION NUMBER: 60/272,409
; PRIOR FILING DATE: 2001-02-28
; PRIOR APPLICATION NUMBER: 60/262,454
; PRIOR FILING DATE: 2001-01-19
; PRIOR APPLICATION NUMBER: 60/276,777
; PRIOR FILING DATE: 2001-03-16
; PRIOR APPLICATION NUMBER: 60/291,672
; PRIOR FILING DATE: 2001-05-17
; PRIOR APPLICATION NUMBER: 60/330,336
; PRIOR FILING DATE: 2001-10-18
; PRIOR APPLICATION NUMBER: 60/265,530
; PRIOR FILING DATE: 2001-01-31
; PRIOR APPLICATION NUMBER: 60/261,376
; PRIOR FILING DATE: 2001-01-16
; NUMBER OF SEQ ID NOS: 269
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 93
; LENGTH: 399
; TYPE: PRT
; ORGANISM: Homo sapiens
US-10-051-874-93

```

Query Match 57.4%; Score 1290; DB 12; Length 399;

```

Best Local Similarity 61.0%; Pred. No. 8.1e-128;
Matches 242; Conservative 59; Mismatches 94; Indels 2; Gaps 1;

Qy 15 MEMLLILVAYMPQNNVNMPTK--AVDPAFNISIIHOQYPCPEYEVATEDGYI 72
Db 1 MKRFLGLVVLWTLHSEGGKLTAVDPETNNVSEIISYWGFPSEYLVETEDGYI 60
Qy 73 LSVNRIPLGLVOPKKTGSRPVVLLQHLGLVGGASNINPNNSLGFILADAGFDVMMGNS 132
Db 61 LCLNRIPLGKXKHSKGPVVPVFLQHLGLADSSNVTNLANSLGFILADAGFDVMMGNS 120
Qy 133 RGNWSRKHTLSIDODEFAFSDYDEMAFDPAPVAVNFILQTKGQEKIYVYGVSOQTMMG 192
Db 121 RGNWSRKHTLSVSDQDEFWAFSDYDEMAKYDLFASINFLNKTGQEQVYVGVSOQTMMG 180
Qy 193 FIAPSTMPPELAOKIKVYFALAPIATVTKHAKSPGTFLPLPDMWIKGLFGKKEFLVQTRL 252
Db 181 FIAPSOIPELAKRIKMPFALGPVAVFCTSPAKLGRPLDHLIKDLFGKKEFLPQSAPL 240
Qy 253 RQVLYLCQGVLDQICSNIMLLGCFNTNNMNGRASVYAAHTLAGTSVQNIHWSQAV 312
Db 241 KWLGTHTVCTHVLKELCGNLCELLCGFNERNLNMSRDVYVTHSPAGTSVQNMHWSQAV 300
Qy 313 NSCELPAFWSGETKNLEKCNQOPTPVVRVVRDMVTPTAMTGGQDWLSNPEDVYMLLSV 372
Db 301 KFORQAFDWSGSAKNYFHNQSYPTTNVKDMLVPTAVWSGHDWLDVYDYNILLTQI 360
Qy 373 TNLVYHKNIPWAHVDFIWLGDAPHMYNEIHLHMQQ 409
Db 361 TNLVFTHEPIWEHLDPIWGLDAPWRLYKNIINLRK 397

RESULT 9
US-10-051-874-94
; Sequence 94, Application US/10051874
; Publication No. US20040005557A1
; GENERAL INFORMATION:
; APPLICANT: Padigar, Muralidhara
; APPLICANT: Alsobrook II, John P
; APPLICANT: Colman, Steven D
; APPLICANT: Spytek, Kimberly A
; APPLICANT: Boldog, Ferenc
; APPLICANT: Vernet, Corine AM
; APPLICANT: Li, Li
; APPLICANT: Shenoy, Suresh G
; APPLICANT: Casman, Stacie J
; APPLICANT: Guo, Xiaojia Sasha
; APPLICANT: Edinger, Shlomit R
; APPLICANT: MacDougall, John R
; APPLICANT: Malyankar, Uriel M
; APPLICANT: Patturajan, Meera
; APPLICANT: Shinkets, Richard A
; APPLICANT: Pena, Carol EA
; APPLICANT: Tchernev, Velizar T
; APPLICANT: Zerhusen, Bryan D
; APPLICANT: Millet, Isabelle
; APPLICANT: Miller, Charles E
; APPLICANT: Lepley, Denise M
; APPLICANT: Smithson, Glenda
; APPLICANT: Baumgartner, Jason C
; APPLICANT: Herrman, John L
; APPLICANT: Peyman, John A
; APPLICANT: Gorman, Linda
; APPLICANT: Mezes, Peter D
; APPLICANT: Kekuda, Ramesh
; APPLICANT: Taupier Jr, Raymond J
; APPLICANT: Gerlach, Valerie
; APPLICANT: Grosse, William M
; APPLICANT: Liu, Xiaohong
; APPLICANT: Ellerman, Karen
; APPLICANT: Rothenberg, Mark
; APPLICANT: Stone, David J
; APPLICANT: Burgess, Catherine E

```

;; TITLE OF INVENTION: PROTEINS, POLYNUCLEOTIDES ENCODING THEM AND METHODS OF
;; TITLE OF INVENTION: USING THE SAME

;; FILE REFERENCE: 21402-245
;; CURRENT APPLICATION NUMBER: US/10/051,874
;; PRIOR FILING DATE: 2002-09-25
;; PRIOR APPLICATION NUMBER: 60/268,595
;; PRIOR FILING DATE: 2001-02-14
;; PRIOR APPLICATION NUMBER: 60/325,306
;; PRIOR FILING DATE: 2001-09-27
;; PRIOR APPLICATION NUMBER: 60/262,587
;; PRIOR FILING DATE: 2001-01-18
;; PRIOR APPLICATION NUMBER: 60/272,409
;; PRIOR FILING DATE: 2001-02-28
;; PRIOR APPLICATION NUMBER: 60/262,454
;; PRIOR FILING DATE: 2001-01-18
;; PRIOR APPLICATION NUMBER: 60/276,777
;; PRIOR FILING DATE: 2001-03-16
;; PRIOR APPLICATION NUMBER: 60/291,672
;; PRIOR FILING DATE: 2001-05-17
;; PRIOR APPLICATION NUMBER: 60/330,336
;; PRIOR FILING DATE: 2001-10-18
;; PRIOR APPLICATION NUMBER: 60/265,530
;; PRIOR FILING DATE: 2001-01-31
;; PRIOR APPLICATION NUMBER: 60/261,376
;; PRIOR FILING DATE: 2001-01-16
;; NUMBER OF SEQ ID NOS: 269
;; SOFTWARE: Patent in Ver. 2.1
;; SEQ ID NO 94
;; LENGTH: 399
;; TYPE: PRT
;; ORGANISM: Homo sapiens
US-10-051-874-94

Query Match 57.4%; Score 1290; DB 12; Length 399;
Best Local Similarity 61.0%; Pred. No. 8.1e-128;
Matches 242; Conservative 59; Mismatches 94; Indels 2; Gaps 1;
QY 15 MEMLLIIVAFQFNVNSVHPTK--AVDPEAFMNIISIIHQGYPCREVEVATEDGYI 72
DB 1 MKGFGLVUCLVLTWTHSEGRKLTAVDPTNMNVSIIISYWGFPSEYLVETEDGYI 60
QY 73 LSVNRIPRGLVOPKKTGRPVVLQHLGVGASNIISLNNPNSLGLFADAGFDVWNGNS 132
DB 61 LCLNRIPHRKXKHSKDGKPKVPVFLQHLGLADSNNWNTLANSSLGLFADAGFDVWNGNS 120
QY 133 RGNWSRKHKTLSDIDQEFWAFSPDEWAPDPAVINFLIKTGOEKIYVGVYSGQTTMG 192
DB 121 RGNWSRKHKTLSDIDQEFWAFSPDEWAPDPAVINFLIKTGOEKIYVGVYSGQTTMG 180
QY 193 FIATFMPPELAQKIMYFALAPIATVYKHAISPGTKFLLPDMNIIKGLFGKKEFLYQTRFL 252
DB 181 FIATFMPPELAQKIMYFALAPIATVYKHAISPGTKFLLPDMNIIKGLFGKKEFLYQTRFL 240
QY 253 RQVLYLCGVTLQDQICSNIMLLGGFTNNMNSRASVYAAHTLAGTSVQNLHWSQAV 312
DB 241 KWLGHVTHVILKELCNLCFLCGFERNLNSRVDVYTTHTSPAGTSVQNLHWSQAV 300
QY 313 NSGELRAFDPWSETKXNLEKCNQPTPVYRVRDMVTPTAMTGGQDMLSNPEDVKMLLSEV 372
DB 301 KFQKFAQFDWSSAKNYPHYNQSYPTTYNVKDLVPTAVWSSGHDVLADVDVYNILLTQI 360
QY 373 TNLVYHKNIPDPAHVDFTWGLDAPHMTNIIHLMQO 409
DB 361 TNLVPHESIPDPAHVDFTWGLDAPHMTNIIHLMQO 397

RESULT 10
US-09-759-130B-411
; Sequence 411, Application US/09759130B
; Publication No. US2003002279A1
; GENERAL INFORMATION:
; APPLICANT: Millennium Pharmaceuticals, Inc.
; APPLICANT: McCarthy, Sean A

;; APPLICANT: Fraser, Christopher C
;; APPLICANT: Sharp, John D
;; APPLICANT: Barnes, Thomas S
;; APPLICANT: Kirt, Susan J
;; APPLICANT: Mackay, Charles R
;; APPLICANT: Myers, Paul S
;; APPLICANT: Leiby, Kevin R
;; APPLICANT: Wrighton, Nicolas
;; APPLICANT: Goodearl, Andrew
;; APPLICANT: Holtzman, Douglas A
;; TITLE OF INVENTION: NOVEL GENES ENCODING PROTEINS HAVING
;; TITLE OF INVENTION: PROGNOSTIC, DIAGNOSTIC, PREVENTIVE, THERAPEUTIC, AND OTHER
;; TITLE OF INVENTION: USES
;; FILE REFERENCE: MPI00-5350NMIM
;; CURRENT APPLICATION NUMBER: US/09/759,130B
;; CURRENT FILING DATE: 2002-09-16
;; PRIOR APPLICATION NUMBER: US 09/479,249
;; PRIOR FILING DATE: 2000-01-07
;; PRIOR APPLICATION NUMBER: US 09/559,497
;; PRIOR FILING DATE: 2000-04-27
;; PRIOR APPLICATION NUMBER: US 09/578,063
;; PRIOR FILING DATE: 2000-05-24
;; PRIOR APPLICATION NUMBER: US 09/333,159
;; PRIOR FILING DATE: 1999-06-14
;; PRIOR APPLICATION NUMBER: US 09/596,194
;; PRIOR FILING DATE: 2000-07-14
;; PRIOR APPLICATION NUMBER: US 09/342,364
;; PRIOR FILING DATE: 1999-06-29
;; PRIOR APPLICATION NUMBER: US 09/608,452
;; PRIOR FILING DATE: 2000-06-30
;; PRIOR APPLICATION NUMBER: US 09/393,996
;; PRIOR FILING DATE: 1999-09-10
;; PRIOR APPLICATION NUMBER: US 09/602,871
;; PRIOR FILING DATE: 2000-06-23
;; PRIOR APPLICATION NUMBER: US 09/420,707
;; PRIOR FILING DATE: 1999-10-19
;; NUMBER OF SEQ ID NOS: 460
;; SOFTWARE: FastSeq for Windows Version 4.0
;; SEQ ID NO 411
;; LENGTH: 399
;; TYPE: PRT
;; ORGANISM: Homo sapiens
US-09-759-130B-411

Query Match 57.4%; Score 1289; DB 11; Length 399;
Best Local Similarity 63.5%; Pred. No. 1e-127;
Matches 235; Conservative 55; Mismatches 80; Indels 0; Gaps 0;
QY 40 AVDPEAFMNIISIIHQGYPCREVEVATEDGYILSVNRIPRGLVOPKKTGRPVVLQHG 99
DB 28 AVDPEAFMNIISIIHQGYPCREVEVATEDGYILSVNRIPRGLVOPKKTGRPVVLQHG 87
QY 100 LVGGASNIISLNNPNSLGLFADAGFDVWNGNSRGNWSRKHKTLSDIDQEFWAFSPYDEM 159
DB 88 LVGGASNIISLNNPNSLGLFADAGFDVWNGNSRGNWSRKHKTLSDIDQEFWAFSPYDEM 147
QY 160 AFEDLPVAVINFLIKTGOEKIYVGVYSGQTTMGFIATFMPPELAQKIMYFALAPIATVK 219
DB 148 AKYDLPASINFLIKTGOEKIYVGVYSGQTTMGFIATFMPPELAQKIMYFALAPIATVK 207
QY 220 HAKSPGTFLLLPDMNIIKGLFGKKEFLYQTRFLQVLYLCGVTLQDQICSNIMLLGGF 279
DB 208 FCTSPMAKLGRLPDHLIKDLFGDKFELPQSAFLKWLGHVTHVILKELCNLCFLCGF 267
QY 280 NTNNMNSRASVYAAHTLAGTSVQNLHWSQAVNSGELRAFDPWSETKXNLEKCNQPTPV 339
DB 268 NERNLNSRVDVYTTHTSPAGTSVQNLHWSQAVNSGELRAFDPWSETKXNLEKCNQPTPV 327
QY 340 YAVRDMVTPTAMTGGQDMLSNPEDVKMLLSEVTLVYHKNIPDPAHVDFTWGLDAPH 399
DB 328 YNVKDLVPTAVWSSGHDVLADVDVYNILLTQITNLVPHESIPDPAHVDFTWGLDAP 387
QY 400 YNEIHLMQO 409

Db 388 YNKIINLWPK 397

RESULT 11

US-10-051-874-92
 ; Sequence 92, Application US/10051874
 ; Publication No. US20040005557A1
 ; GENERAL INFORMATION:
 ; APPLICANT: Padigaru, Muralidhara
 ; APPLICANT: Alsobrook II, John P
 ; APPLICANT: Colman, Steven D
 ; APPLICANT: Spytek, Kimberly A
 ; APPLICANT: Boldog, Ferenc
 ; APPLICANT: Vernet, Corine AM
 ; APPLICANT: Li, Li
 ; APPLICANT: Shenoy, Suresh G
 ; APPLICANT: Casman, Stacie J
 ; APPLICANT: Guo, Xiaojia Sashu
 ; APPLICANT: Edinger, Shlomit R
 ; APPLICANT: MacDougall, John R
 ; APPLICANT: Malyankar, Uziel M
 ; APPLICANT: Patturajan, Meera
 ; APPLICANT: Shankets, Richard A
 ; APPLICANT: Pena, Carol EA
 ; APPLICANT: Tchernev, Velizar T
 ; APPLICANT: Zerhusen, Bryan D
 ; APPLICANT: Millet, Isabelle
 ; APPLICANT: Miller, Charles E
 ; APPLICANT: Leplev, Denise M
 ; APPLICANT: Smithson, Glennda
 ; APPLICANT: Baumgartner, Jason C
 ; APPLICANT: Peyman, John L
 ; APPLICANT: Gorman, Linda
 ; APPLICANT: Mezes, Peter D
 ; APPLICANT: Kekuda, Ramesh
 ; APPLICANT: Taupier Jr, Raymond J
 ; APPLICANT: Gerlach, Valerie
 ; APPLICANT: Grosse, William M
 ; APPLICANT: Liu, Xiaohong
 ; APPLICANT: Ellerman, Karen
 ; APPLICANT: Rothenberg, Mark
 ; APPLICANT: Stone, David J
 ; APPLICANT: Burgess, Catherine E
 ; TITLE OF INVENTION: PROTEINS, POLYNUCLEOTIDES ENCODING THEM AND METHODS OF
 ; FILE REFERENCE: 21402-245
 ; CURRENT APPLICATION NUMBER: US/10/051,874
 ; PRIOR FILING DATE: 2002-09-25
 ; PRIOR APPLICATION NUMBER: 60/268,595
 ; PRIOR FILING DATE: 2001-02-14
 ; PRIOR APPLICATION NUMBER: 60/325,306
 ; PRIOR FILING DATE: 2001-09-27
 ; PRIOR APPLICATION NUMBER: 60/262,587
 ; PRIOR FILING DATE: 2001-01-18
 ; PRIOR APPLICATION NUMBER: 60/272,409
 ; PRIOR FILING DATE: 2001-02-28
 ; PRIOR APPLICATION NUMBER: 60/262,454
 ; PRIOR FILING DATE: 2001-01-18
 ; PRIOR APPLICATION NUMBER: 60/276,777
 ; PRIOR FILING DATE: 2001-03-16
 ; PRIOR APPLICATION NUMBER: 60/291,672
 ; PRIOR FILING DATE: 2001-05-17
 ; PRIOR APPLICATION NUMBER: 60/330,336
 ; PRIOR FILING DATE: 2001-10-18
 ; PRIOR APPLICATION NUMBER: 60/265,530
 ; PRIOR FILING DATE: 2001-01-31
 ; PRIOR APPLICATION NUMBER: 60/261,376
 ; PRIOR FILING DATE: 2001-01-16
 ; NUMBER OF SEQ ID NOS: 269
 ; SOFTWARE: PatentIn Ver. 2.1
 ; SEQ ID NO 92

; LENGTH: 399
 ; TYPE: PRT
 ; ORGANISM: Homo sapiens
 US-10-051-874-92

Query Match 57.4%; Score 1289; DB 12; Length 399;
 Best Local Similarity 63.5%; Pred. No. 1e-127;
 Matches 235; Conservative 55; Mismatches 80; Indels 0; Gaps 0;

QY 40 AVDPFAFNWISIIHQGYPCPEYEVEDGCVILSVNRIPRGLVQPKTKGSRPVVLLQHG 99
 DB 28 AVDPETNNVSEIISYWGFPSEYLVETEDGYILCLNRPGRKXNSDKGPKPVVFLQHG 87
 QY 100 LVGGASNWSLNPNNISLGFILADAGFDVWNGSRGNSRKNHKTLSIDODEFWAFSYDEM 159
 DB 88 LLDSSNWNVLNLSLGFILADAGFDVWNGSRGNTWSRKHKTLSVSDQDFWAFSYDEM 147
 QY 160 ARFDLPVAVINFILQTKGQEKIYYVGYSGQTTMGFTAFSTMPDLAQIKWYFALAPIATVK 219
 DB 148 KYVDLPASINFILNKTGQEQVYVVGHSQTTIGTFIAFSQIPELAKRIOMFFALGPVASVA 207
 QY 220 HAKSPGTAKFLLPDMWIKGLPKGKFLYQTRFLQVLYLGGQVILDOICSNIMLLGCP 279
 DB 208 FCTSPWAKLGRPLDHLIKDLFDKKEFLQPSAFLKWLGHVCTHVLKELCGNLCFLGCP 267
 QY 280 NTNNMNSRASVYAAHTLAGTSVONILHWSQAVNSGELRAFDMGSETKNLEKCKNQTPVR 339
 DB 268 NEENLMSRVDTYTHSPAGTSVQNMHWSQAVKQFQAFDMGSSAKNYHYNQSYPT 327
 QY 340 YVRDMTPTMTGQGDWLNPNEDVXMLLSEVNLIVHKNIPBWAHVDFTWGLDAPHRM 399
 DB 328 YNVKDLVPTAVWSGHDWLDVVDVNNILLQTITLVPHESIPEWEHLDFIWLGDAPWRL 387
 QY 400 YNEIITHLMQO 409
 DB 388 YNKIINLWPK 397

RESULT 12

US-10-042-431-41
 ; Sequence 41, Application US/10042431
 ; Publication No. US20020182675A1
 ; GENERAL INFORMATION:
 ; APPLICANT: MCCARTHY, Sean A
 ; APPLICANT: BARNES, Thomas M
 ; APPLICANT: FRASER, Christopher C
 ; APPLICANT: SHARP, John D
 ; TITLE OF INVENTION: NOVEL GENES ENCODING PROTEINS HAVING DIAGNOSTIC,
 ; FILE REFERENCE: 10147-6U2
 ; CURRENT APPLICATION NUMBER: US/10/042,431
 ; CURRENT FILING DATE: 2001-10-25
 ; PRIOR APPLICATION NUMBER: US 09/333,159
 ; PRIOR FILING DATE: 1999-06-14
 ; PRIOR APPLICATION NUMBER: US 09/578,063
 ; PRIOR FILING DATE: 2000-05-24
 ; NUMBER OF SEQ ID NOS: 79
 ; SOFTWARE: PatentIn Ver. 2.1
 ; SEQ ID NO 41
 ; LENGTH: 399
 ; TYPE: PRT
 ; ORGANISM: Homo sapiens
 US-10-042-431-41

Query Match 57.4%; Score 1289; DB 14; Length 399;
 Best Local Similarity 63.5%; Pred. No. 1e-127;
 Matches 235; Conservative 55; Mismatches 80; Indels 0; Gaps 0;

QY 40 AVDPFAFNWISIIHQGYPCPEYEVEDGCVILSVNRIPRGLVQPKTKGSRPVVLLQHG 99
 DB 28 AVDPETNNVSEIISYWGFPSEYLVETEDGYILCLNRPGRKXNSDKGPKPVVFLQHG 87
 QY 100 LVGGASNWSLNPNNISLGFILADAGFDVWNGSRGNSRKNHKTLSIDODEFWAFSYDEM 159

```

Db      88  LADSNWNTNLSNLSGLFADAGFDVWNGSRGNTSRKHTLSVSDQDFWAFSYDEM 147
QY      160  ARPDLPAVINFILOKTOGQKIYVYVSGTGTGFIATSTMPBLAQKIRYFALAPIATVK 219
Db      148  AKYDLPAVINFILOKTOGQKIYVYVSGTGTGFIATSTMPBLAQKIRYFALAPIATVK 207
QY      220  HAKSPGTEKLLPDMWIKGLFGKKEFLYOTRFLQVILYVLCQVILDOICSNIMLLGCF 279
Db      208  FCTSPWAKLGRUPDLHILKDLGKFLPOSALFKLWGHVCHVILKELCNLCFLGCF 267
QY      280  NNTNMMSRASYAHTLAGTSVQNLHWSQAVNSGELRAFDWGSSETKXLEKCNQPTVR 339
Db      268  NERNLANSRVDVYTHSPAGTSVQNLHWSQAVNSGELRAFDWGSSETKXLEKCNQPTVR 327
QY      340  YVRDVTMTVTAQKQDMLSPEDVYVNLVHKNIPWAHVDFFIWLGDAPHRM 399
Db      328  YVNDMLVPTAVWSGGHDLADVDVNNILLTQITNLVPHESIPWEHLDFFIWLGDAPWRL 387
QY      400  YNEIHLMOQ 409
Db      388  YNKIINLMK 397

```

```

RESULT 13
US-10-003-302-4
; Sequence 4, Application US/10003302
; Publication No. US20020142435A1
; GENERAL INFORMATION:
; APPLICANT: MERKULOV, Gennady et al.
; TITLE OF INVENTION: ISOLATED HUMAN LIPASE PROTEINS, NUCLEIC
; TITLE OF INVENTION: ACID MOLECULES ENCODING HUMAN LIPASE PROTEINS, AND USES
; TITLE OF INVENTION: THEREOF
; FILE REFERENCE: CL001186DIV
; CURRENT APPLICATION NUMBER: US/10/003,302
; CURRENT FILING DATE: 2001-12-06
; NUMBER OF SEQ ID NOS: 4
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 4
; LENGTH: 392
; TYPE: PRT
; ORGANISM: Homo sapiens
US-10-003-302-4

```

```

Query Match      57.1%; Score 1282; DB 14; Length 392;
Best Local Similarity 63.8%; Pred. No. 5.5e-127;
Matches 234; Conservative 53; Mismatches 80; Indels 0; Gaps 0;

QY      40  AVDPETNNVSEIIHQGYPCPEYEVATEDGYILSVNRPRLGVLQPKTGRPVVLLQHG 99
Db      26  AVDPETNNVSEIIHQGYPCPEYEVATEDGYILSVNRPRLGVLQPKTGRPVVLLQHG 85
QY      100  LVGGSANWISNLPNLSGLFADAGFDVWNGSRGNTSRKHTLSVSDQDFWAFSYDEM 159
Db      86  LADSNWNTNLSNLSGLFADAGFDVWNGSRGNTSRKHTLSVSDQDFWAFSYDEM 145
QY      160  ARPDLPAVINFILOKTOGQKIYVYVSGTGTGFIATSTMPBLAQKIRYFALAPIATVK 219
Db      146  AKYDLPAVINFILOKTOGQKIYVYVSGTGTGFIATSTMPBLAQKIRYFALAPIATVK 205
QY      220  HAKSPGTEKLLPDMWIKGLFGKKEFLYOTRFLQVILYVLCQVILDOICSNIMLLGCF 279
Db      206  FCTSPWAKLGRUPDLHILKDLGKFLPOSALFKLWGHVCHVILKELCNLCFLGCF 265
QY      280  NNTNMMSRASYAHTLAGTSVQNLHWSQAVNSGELRAFDWGSSETKXLEKCNQPTVR 339
Db      266  NERNLANSRVDVYTHSPAGTSVQNLHWSQAVNSGELRAFDWGSSETKXLEKCNQPTVR 325
QY      340  YVRDVTMTVTAQKQDMLSPEDVYVNLVHKNIPWAHVDFFIWLGDAPHRM 399
Db      326  YVNDMLVPTAVWSGGHDLADVDVNNILLTQITNLVPHESIPWEHLDFFIWLGDAPWRL 385
QY      400  YNEIHL 406

```

```

Db      386  YNKIINL 392
RESULT 14
US-09-759-130B-420
; Sequence 420, Application US/09759130B
; Publication No. US20030022279A1
; GENERAL INFORMATION:
; APPLICANT: Millennium Pharmaceuticals, Inc.
; APPLICANT: McCarthy, Sean A
; APPLICANT: Fraser, Christopher C
; APPLICANT: Sharp, John D
; APPLICANT: Barnes, Thomas S
; APPLICANT: Kirt, Susan J
; APPLICANT: Mackay, Charles R
; APPLICANT: Myers, Paul S
; APPLICANT: Leiby, Kevin R
; APPLICANT: Wrighton, Nicolas
; APPLICANT: Goodearl, Andrew
; APPLICANT: Holtzman, Douglas A
; TITLE OF INVENTION: NOVEL GENES ENCODING PROTEINS HAVING
; TITLE OF INVENTION: PROGNOSTIC, DIAGNOSTIC, PREVENTIVE, THERAPEUTIC, AND OTHER
; TITLE OF INVENTION: USES.
; FILE REFERENCE: MP100-5350NMIM
; CURRENT APPLICATION NUMBER: US/09/759,130B
; CURRENT FILING DATE: 2002-09-16
; PRIOR APPLICATION NUMBER: US 09/479,249
; PRIOR FILING DATE: 2000-01-07
; PRIOR APPLICATION NUMBER: US 09/559,497
; PRIOR FILING DATE: 2000-04-27
; PRIOR APPLICATION NUMBER: US 09/578,063
; PRIOR FILING DATE: 2000-05-24
; PRIOR APPLICATION NUMBER: US 09/333,159
; PRIOR FILING DATE: 1999-06-14
; PRIOR APPLICATION NUMBER: US 09/596,194
; PRIOR FILING DATE: 2000-07-14
; PRIOR APPLICATION NUMBER: US 09/342,364
; PRIOR FILING DATE: 1999-06-29
; PRIOR APPLICATION NUMBER: US 09/608,452
; PRIOR FILING DATE: 2000-06-30
; PRIOR APPLICATION NUMBER: US 09/393,996
; PRIOR FILING DATE: 1999-09-10
; PRIOR APPLICATION NUMBER: US 09/602,871
; PRIOR FILING DATE: 2000-06-23
; PRIOR APPLICATION NUMBER: US 09/420,707
; PRIOR FILING DATE: 1999-10-19
; NUMBER OF SEQ ID NOS: 460
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 420
; LENGTH: 221
; TYPE: PRT
; ORGANISM: Homo sapiens
US-09-759-130B-420

```

```

Query Match      51.8%; Score 1165; DB 11; Length 221;
Best Local Similarity 100.0%; Pred. No. 6e-115;
Matches 221; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY      34  VHMPTKAVDPEAPNLSIIHQGYPCPEYEVATEDGYILSVNRPRLGVLQPKTGRSV 93
Db      1  VHMPTKAVDPEAPNLSIIHQGYPCPEYEVATEDGYILSVNRPRLGVLQPKTGRSV 60
QY      94  VLLQHLGVGASNWSNLPNLSGLFADAGFDVWNGSRGNTSRKHTLSVSDQDFW 153
Db      61  VLLQHLGVGASNWSNLPNLSGLFADAGFDVWNGSRGNTSRKHTLSVSDQDFW 120
QY      154  FSYDEWARFDLPVAFILQKTOGQKIYVYVSGTGTGFIATSTMPBLAQKIRYFAL 213
Db      121  FSYDEWARFDLPVAFILQKTOGQKIYVYVSGTGTGFIATSTMPBLAQKIRYFAL 180
QY      214  PIATVHKASPGTKFLLLPDMIKGLFGKKEFLYOTRFLRQ 254

```

Db 181 PIATVKHAKSPGTFKLLPDDMMIKGLFGKKEFLYOTRFLRQ 221

RESULT 15

US-10-042-431-50
 ; Sequence 50, Application US/10042431
 ; Publication No. US20020182675A1
 ; GENERAL INFORMATION:
 ; APPLICANT: MCCARTHY, Sean A
 ; APPLICANT: BARNES, Thomas M
 ; APPLICANT: FRASER, Christopher C
 ; APPLICANT: SHARP, John D
 ; TITLE OF INVENTION: NOVEL GENES ENCODING PROTEINS HAVING DIAGNOSTIC,
 ; FILE REFERENCE: 10147-GU2
 ; CURRENT APPLICATION NUMBER: US/10/042,431
 ; PRIOR FILING DATE: 2001-10-25
 ; PRIOR APPLICATION NUMBER: US 09/333,159
 ; PRIOR FILING DATE: 1999-06-14
 ; PRIOR APPLICATION NUMBER: US 09/578,063
 ; PRIOR FILING DATE: 2000-05-24
 ; NUMBER OF SEQ ID NOS: 79
 ; SOFTWARE: PatentIn Ver. 2.1
 ; SEQ ID NO 50
 ; LENGTH: 221
 ; TYPE: PRT
 ; ORGANISM: Homo sapiens
 US-10-042-431-50

Query Match 51.8%; Score 1165; DB 14; Length 221;
 Best Local Similarity 100.0%; Pred. No. 6e-115;
 Matches 221; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 QY 34 VHMPTKAVDPEAFMNISEIIHQGYPCPEYEVATEDGYILSVNRIPRGLVQPKKTGSRPV 93
 DB 1 VHMPTKAVDPEAFMNISEIIHQGYPCPEYEVATEDGYILSVNRIPRGLVQPKKTGSRPV 60
 QY 94 VLLQHLVCGAGNWSINLNNLSGLFILADAGPDVWNGSRGNWGRKHKHTLSIDQDEFWA 153
 DB 61 VLLQHLVCGAGNWSINLNNLSGLFILADAGPDVWNGSRGNWGRKHKHTLSIDQDEFWA 120
 QY 154 FSYDEMARFDLPAVINFILOKTQGEKIYVGVYSGQTTWGFIAPFTMPPELAKIKMYFALA 213
 DB 121 FSYDEMARFDLPAVINFILOKTQGEKIYVGVYSGQTTWGFIAPFTMPPELAKIKMYFALA 180
 QY 214 PIATVKHAKSPGTFKLLPDDMMIKGLFGKKEFLYOTRFLRQ 254
 DB 181 PIATVKHAKSPGTFKLLPDDMMIKGLFGKKEFLYOTRFLRQ 221

Search completed: February 19, 2004, 14:34:24
 Job time : 30.6162 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2004 Compugen Ltd.

OM protein - protein search, using sw model

Run on: February 19, 2004, 14:19:24 ; Search time 14.5683 Seconds
(without alignments)
1228.526 Million cell updates/sec

Title: US-10-042-431-47
Perfect score: 2247
Sequence: 1 MLETLRQWIVSHRMENMLL.....IHLMQQETNLSQRCCEAVL 423

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 328717 seqs, 42310858 residues

Total number of hits satisfying chosen parameters: 328717

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : Issued Patents AA.*
1: /cgn2_6/ptodata/2/iaa/5A_COMB.pep.*
2: /cgn2_6/ptodata/2/iaa/5B_COMB.pep.*
3: /cgn2_6/ptodata/2/iaa/6A_COMB.pep.*
4: /cgn2_6/ptodata/2/iaa/6B_COMB.pep.*
5: /cgn2_6/ptodata/2/iaa/PTUS_COMB.pep.*
6: /cgn2_6/ptodata/2/iaa/backfiles1.pep.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	1282	57.1	392	4	US-09-820-001-4
2	1161	51.7	377	1	US-08-227-108-17
3	1161	51.7	377	2	US-09-073-674-17
4	1150.5	51.2	379	1	US-08-227-108-18
5	1150.5	51.2	379	2	US-09-073-674-18
6	1141.5	50.8	375	4	US-09-348-930A-6
7	1141.5	50.8	379	1	US-08-227-108-3
8	1141.5	50.8	379	4	US-09-348-930A-2
9	1141.5	50.8	380	1	US-08-227-108-5
10	1141.5	50.8	380	2	US-09-073-674-5
11	1141.5	50.8	380	1	US-09-185-489-2
12	1127	50.2	378	4	US-10-043-655B-2
13	1119	49.8	398	4	US-09-348-930A-9
14	1098.5	48.9	380	1	US-08-227-108-16
15	1098.5	48.9	380	2	US-09-073-674-16
16	1098.5	48.9	380	1	US-09-820-001-2
17	1031	45.9	395	4	US-09-348-930A-4
18	1003	44.6	325	4	US-09-348-930A-30091
19	143.5	6.4	348	4	US-08-484-105-22
20	105	4.7	430	1	US-08-484-106-22
21	105	4.7	430	1	US-08-484-106-22
22	103	4.6	292	4	US-09-328-352-7538
23	102	4.5	346	2	US-08-602-359A-34
24	101.5	4.5	312	4	US-09-107-532A-5684
25	98.5	4.4	388	1	US-08-232-519-2
26	98.5	4.4	388	1	US-08-456-956-2
27	96.5	4.3	287	4	US-09-584-568C-4

ALIGNMENTS

RESULT 1

US-09-820-001-4
; Sequence 4, Application US/09820001
; Patent No. 6387680
; GENERAL INFORMATION:
; APPLICANT: MERKULOV, Gennady et al
; TITLE OF INVENTION: ISOLATED HUMAN LIPASE PROTEINS, NUCLEIC
; TITLE OF INVENTION: ACID MOLECULES ENCODING HUMAN LIPASE PROTEINS, AND USES
; FILE REFERENCE: CL001186
; CURRENT APPLICATION NUMBER: US/09/820,001
; CURRENT FILING DATE: 2001-03-29
; NUMBER OF SEQ ID NOS: 4
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 4
; LENGTH: 392
; TYPE: PRT
; ORGANISM: Human
US-09-820-001-4

Query Match 57.1%; Score 1282; DB 4; Length 392;
Best Local Similarity 63.6%; Pred. No. 6.5e-127;
Matches 234; Conservative 53; Mismatches 80; Indels 0; Gaps 0;

QY	40	AVDPEAFNMISRIHQGYPCPEEYEVATEDGYILSVNRIPIRGLVQPKTKGSRPVVLQHG	99
DB	26	AVDPETNMNVSEIIISYWGFPSEYIVETEDGYILCLNRIPIHGRKNHSDKGPVVFVLQHG	85
QY	100	LVGGASNWTISNLPNNLSGLFILADAGFDYWGNSRGNANRKHKTLSIDODEFWAFSYDEM	159
DB	86	LIADSSNWTNLSNLSGLFILADAGFDYWGNSRGNANRKHKTLSIDODEFWAFSYDEM	145
QY	160	ARFDPFAVINFLQTKGQEKIYVGYSGQTTMGFTAFSTMPBLAQIKMYFALAPIATVK	219
DB	146	AKYDLPASINFILNKTGQEVVYVGHSGQTIGTIGTAFSQIPELAKIRKFPALGPVASVA	205
QY	220	HAKSGPTKELLIPDMMIKGLFKKFLYQTRFLRQVLYLCSQVILDOICSNIMLLCGP	279
DB	206	FCTSPMAKLGRIELIKDLFGDKFLPQSAFLKWLGHVCTHVTILKELCGNLCLCGP	265
QY	280	NTNNMMSRASVYAHTLAGTSVQNLHWSQVNSGELRAPDWGSETKNLEKCNQPTVR	339
DB	266	NERNLNMSRDVYTHSPAGTSVQNLHWSQVNSGELRAPDWGSETKNLEKCNQPTVR	325
QY	340	YVRDMVTFTAMWTGGQDMLSNPDKMLLSEVTLNLIYHKNIPENAHVDFIWLGDAPHRM	399
DB	326	YXVKDMLVPTAVNSGHDWLDVVDVNIILLTQITNLVFHESIPWEHLDFIWGLDAPWRL	385
QY	400	YNEIHL 406	

[illegible]

```

RESULT 4
US-08-227-108-18
; Sequence 18, Application US/08227108
; Patent No. 5807726
; GENERAL INFORMATION:
; APPLICANT: Blanchard, Claude
; APPLICANT: Benichourt, Claude
; APPLICANT: Junien, Jean-Louis
; TITLE OF INVENTION: Recombinant Dog Gastric Lipase
; NUMBER OF SEQUENCES: 21
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Pennie & Edmonds
; STREET: 1155 Avenue of the Americas
; CITY: New York
; STATE: New York
; COUNTRY: U.S.A.
; ZIP: 10036
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/227,108
; FILING DATE: 03-APR-1994
; CLASSIFICATION: 435

```

Query Match	51.2%	Score 1150.5	DB 1	Length 379
Best Local Similarity	56.8%	Pred. No. 4.9e-113		
Matches 213	Conservative 65	Mismatches 94	Indels 3	Gaps 1
QY	37	PTKAVDPFAFMNISIIIOHQYCPCEEVEATGDIYLLSVNRIPRGVLQPKKTGSRPVLL	96	
Db	7	PT--NPEVNNNISOMISYGYPSKKEVVTGDIYLVNRRIPYKCKNSGRGQPPVFL	63	
QY	97	QHLGVGASNIWSNLPNNSLGFILADAGFDVWNGSNGRGNARSKHKTLLSIDODFWARSY	156	
Db	64	QHLGLASASNIWSNLPNNSLAFILADAGYGVNLGSRGNTWSRNLYYSPDSVBEWAFSE	123	
QY	157	DEMARFDPALVINFILOKTGQEKLYYGVYSGITMTGFIAPSTMPELAOKIKMYFALAPIA	216	
Db	124	DEMAKYDLPATIDFIVKETGQEKLVYGVHSGQITIGFIAPSTNPKLAERIKTFYALAPVA	183	
QY	217	TVYKHAKSPTGTEKLLLPDMNMIKLGFEKKEFIYQTRFRLQLVYLCGVQILDQCSNIMLLL	276	
Db	184	TVYKTKSLVNKLRFIPPTMFKIIIFGDKIFYPHNPFDQFLATQVCSRETLLVCSKALPFI	243	
QY	277	GGFTNNNNMGSRASYAHTLAGTSVQNIILHWSQAVNSGELRAFDWGSSTKLEKCNQPT	336	

Db	244	CGFDSANLMSRLDYYVSHNPAGTSVONMLHWTQAVKSGNFQAFNWGSPAQNVVHFNQPT	303
Qy	337	PVRYKVRDWTPTAMWTCGGQDWLSNPBDYKLLSEVNLIIYHKNIPEWAHVDIFWGLDAP	396
Db	304	PFYQVNTAMVPFLAVWSGNDWLADPDQVDVLLLPKLSNLIYHKEILFYPYHLODFINAMNAP	363
Qy	397	HRMYNEIHLMOQEE	411
Db	364	QEVYNEIISMMAKDK	378

```

1
2
3
4
5
6
7
8
9
10
11
12
13
14
15
16
17
18
19
20
21
22
23
24
25
26
27
28
29
30
31
32
33
34
35
36
37
38
39
40
41
42
43
44
45
46
47
48
49
50
51
52
53
54
55
56
57
58
59
60
61
62
63
64
65
66
67
68
69
70
71
72
73
74
75
76
77
78
79
80
81
82
83
84
85
86
87
88
89
90
91
92
93
94
95
96
97
98
99
100
101
102
103
104
105
106
107
108
109
110
111
112
113
114
115
116
117
118
119
120
121
122
123
124
125
126
127
128
129
130
131
132
133
134
135
136
137
138
139
140
141
142
143
144
145
146
147
148
149
150
151
152
153
154
155
156
157
158
159
160
161
162
163
164
165
166
167
168
169
170
171
172
173
174
175
176
177
178
179
180
181
182
183
184
185
186
187
188
189
190
191
192
193
194
195
196
197
198
199
200
201
202
203
204
205
206
207
208
209
210
211
212
213
214
215
216
217
218
219
220
221
222
223
224
225
226
227
228
229
230
231
232
233
234
235
236
237
238
239
240
241
242
243
244
245
246
247
248
249
250
251
252
253
254
255
256
257
258
259
260
261
262
263
264
265
266
267
268
269
270
271
272
273
274
275
276
277
278
279
280
281
282
283
284
285
286
287
288
289
290
291
292
293
294
295
296
297
298
299
300
301
302
303
304
305
306
307
308
309
310
311
312
313
314
315
316
317
318
319
320
321
322
323
324
325
326
327
328
329
330
331
332
333
334
335
336
337
338
339
340
341
342
343
344
345
346
347
348
349
350
351
352
353
354
355
356
357
358
359
360
361
362
363
364
365
366
367
368
369
370
371
372
373
374
375
376
377
378
379
380
381
382
383
384
385
386
387
388
389
390
391
392
393
394
395
396
397
398
399
400
401
402
403
404
405
406
407
408
409
410
411
412
413
414
415
416
417
418
419
420
421
422
423
424
425
426
427
428
429
430
431
432
433
434
435
436
437
438
439
440
441
442
443
444
445
446
447
448
449
450
451
452
453
454
455
456
457
458
459
460
461
462
463
464
465
466
467
468
469
470
471
472
473
474
475
476
477
478
479
480
481
482
483
484
485
486
487
488
489
490
491
492
493
494
495
496
497
498
499
500
501
502
503
504
505
506
507
508
509
510
511
512
513
514
515
516
517
518
519
520
521
522
523
524
525
526
527
528
529
530
531
532
533
534
535
536
537
538
539
540
541
542
543
544
545
546
547
548
549
550
551
552
553
554
555
556
557
558
559
560
561
562
563
564
565
566
567
568
569
570
571
572
573
574
575
576
577
578
579
580
581
582
583
584
585
586
587
588
589
590
591
592
593
594
595
596
597
598
599
600
601
602
603
604
605
606
607
608
609
610
611
612
613
614
615
616
617
618
619
620
621
622
623
624
625
626
627
628
629
630
631
632
633
634
635
636
637
638
639
640
641
642
643
644
645
646
647
648
649
650
651
652
653
654
655
656
657
658
659
660
661
662
663
664
665
666
667
668
669
670
671
672
673
674
675
676
677
678
679
680
681
682
683
684
685
686
687
688
689
690
691
692
693
694
695
696
697
698
699
700
701
702
703
704
705
706
707
708
709
710
711
712
713
714
715
716
717
718
719
720
721
722
723
724
725
726
727
728
729
730
731
732
733
734
735
736
737
738
739
740
741
742
743
744
745
746
747
748
749
750
751
752
753
754
755
756
757
758
759
760
761
762
763
764
765
766
767
768
769
770
771
772
773
774
775
776
777
778
779
780
781
782
783
784
785
786
787
788
789
790
791
792
793
794
795
796
797
798
799
800
801
802
803
804
805
806
807
808
809
810
811
812
813
814
815
816
817
818
819
820
821
822
823
824
825
826
827
828
829
830
831
832
833
834
835
836
837
838
839
840
841
842
843
844
845
846
847
848
849
850
851
852
853
854
855
856
857
858
859
860
861
862
863
864
865
866
867
868
869
870
871
872
873
874
875
876
877
878
879
880
881
882
883
884
885
886
887
888
889
890
891
892
893
894
895
896
897
898
899
900
901
902
903
904
905
906
907
908
909
910
911
912
913
914
915
916
917
918
919
920
921
922
923
924
925
926
927
928
929
930
931
932
933
934
935
936
937
938
939
940
941
942
943
944
945
946
947
948
949
950
951
952
953
954
955
956
957
958
959
960
961
962
963
964
965
966
967
968
969
970
971
972
973
974
975
976
977
978
979
980
981
982
983
984
985
986
987
988
989
990
991
992
993
994
995
996
997
998
999

```

	Query Match	51.2%	Score 1150.5	DB 2	Length 379
	Best Local Similarity	56.8%	Pred. No. 4.9e-113		
	Matches 213	Conservative 65	Mismatches 94	Indels 3	Gaps 1
Qy	37	PTRAVDPEAPMNISEIIHQGYPCBEYEYVATEDGYILSVNRIRPGLVQPKTKTSRPPVLL	96		
Db	7	PT---NPEVMNINISQMSIYWGVPSEKYEVVTEDEGYILEVNRIRPYGKKNSSNRCQRPVPL	63		
Qy	97	QHGLVGGASNIWLNPNNSIGFTLADAGPDVVMGNSRGNASWKSKHTLTSLDQDEFWAFSY	156		
Db	64	QHGLLASASNIWLNPNNSLAFLLADAGYVWLGNRSGNWTNRNLIYYSPDSVEFFWAFSF	123		
Qy	157	DEWARDLPAVINFILOKQGEKIYYVVSQGTMTGFIAPSTWPELAQKIKMTYALAPYA	216		
Db	124	DEMAKYDLPATIDFIVKSTGQEKLVHVGHSQGTTFIGFIAPSTNPVKLAERIKTFYALAPYA	183		
Qy	217	TVVHKASPGPKFLLLPDMMIKGLFGKKEFLYQTRFLRQLVLYLCGVVILDQICSNLMILL	276		

Db 184 TVXYTKSLVNLKRLFPPTMPKLIIPGDKIYPFHNFFDQFLATQVCSRETILNVICSNALPFI 243
Qy 277 GGFNTNMMNSRASVYAAHTLACTSVQNTILHWSQAVNSGELRAFDPWGSSTKLEKCNQPT 336
Db 244 CGFDSANLNSRLDVTYVSHNPAGTSVQNLHWTQAVKSGNFQAFNWSGPAQNVHFNQPT 303
Qy 337 PVRYRVDMTVPTAMTGGQDMLSNPDEVKMLSEVTNLIYHKNIPDEWAHVDPIWGLDAP 396
Db 304 PPYYNTAMNVPVAVWSGGNDLADPDQVDLLLPKLSNLIYHKEILPYNHLDFIWMANAP 363
Qy 397 HRMNEIHLMOQEE 411
Db 364 QEVYNEIISMAKDK 378

RESULT 6
US-09-348-930A-6
; Sequence 6, Application US/09348930A
; Patent No. 6573431
; GENERAL INFORMATION:
; APPLICANT: Lenee, P
; APPLICANT: Gruber, V
; APPLICANT: Baudino, S
; APPLICANT: Merot, B
; APPLICANT: Benicourt, C
; TITLE OF INVENTION: Recombinant Preducenal Lipases and Polypeptides Derivatives Prod
; TITLE OF INVENTION: Plants, Processes for Obtaining Them and Their Uses
; FILE REFERENCE: 18433/2012
; CURRENT APPLICATION NUMBER: US/09/348,930A
; CURRENT FILING DATE: 1999-07-02
; PRIOR APPLICATION NUMBER: 08/945,321
; PRIOR FILING DATE: 1998-02-12
; PRIOR APPLICATION NUMBER: PCT/FR96/00606
; PRIOR FILING DATE: 1996-03-19
; NUMBER OF SEQ ID NOS: 37
; SOFTWARE: PatentIn version 3.0
; SEQ ID NO 6
; LENGTH: 375
; TYPE: PRT
; ORGANISM: Canis familiaris
US-09-348-930A-6

Query Match 50.8%; Score 1141.5; DB 4; Length 375;
Best Local Similarity 56.5%; Pred. No. 4.3e-112;
Matches 210; Conservative 67; Mismatches 90; Indels 5; Gaps 3;

Qy 37 PTKAVDPEAFNISEIIHQGYPCPEYEVEATEDGYILSVNRIIPRGLVQPKKTSRPPVLL 96
Db 3 PT---NPEVTNISIOMITYMGYPABEYEVVTEDEGYILGIDRIPYGRKNSENIGRRPVAFL 59
Qy 97 QHGLVGGASNWSNLPPNSLGFILADAGFDVWMCNSRGNWSKHKHTLSIDQDEFWAFSY 156
Db 60 QHGLASATNWSNLPPNSLAFILADAGYDVWLCNSRGNTWARNLNLYSPDSVEFWAFSF 119
Qy 157 DEMARFDPVAVINFILQKTEKIIYVYSGTGTMGFIAPSTWPELAQKIMYFALAPIA 216
Db 120 DEMAKYDLPATIDFLLKTKTGDKLHYVGHSGTGTIGFIAPSTWPELAQKIMYFALAPIA 179
Qy 217 TVKHAQSPGTFKLLPDMMLKGLFGKKEFLYQTRFLQ-LVIYLCGVQLDQCSNIMLL 275
Db 180 TVKTEILLNKLMLVPFSLFKLIFGNKIP-YPHHFFDQFLATEVCSRETVDLLCSNALFI 238
Qy 276 LGGFNTNMMNSRASVYAAHTLACTSVQNTILHWSQAVNSGELRAFDPWGSSTKLEKCNQPT 335
Db 239 ICGFDTNMLNSRLDVTYVSHNPAGTSVQNLHWSQAVKSGKQAFDWSGVPQNMHHQS 298
Qy 336 TPVRYRVDMTVPTAMTGGQDMLSNPDEVKMLSEVTNLIYHKNIPDEWAHVDPIWGLDA 395
Db 299 MPYYNTLDMHVPVAVWSGGNDLADPHDVLNLSKLPNLIYHKKIPPPYHNLDFIWMADA 358
Qy 396 PHRYNEIHL 407

Db 359 PQAYNEIVSNM 370
RESULT 7
US-08-227-108-3
; Sequence 3, Application US/08227108
; Patent No. 5807726
; GENERAL INFORMATION:
; APPLICANT: Blanchard, Claire
; APPLICANT: Benicourt, Claude
; APPLICANT: Junien, Jean-Louis
; TITLE OF INVENTION: Recombinant Dog Gastric Lipase
; NUMBER OF SEQUENCES: 21
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Pennie & Edmonds
; STREET: 1155 Avenue of the Americas
; CITY: New York
; STATE: New York
; COUNTRY: U.S.A.
; ZIP: 10036
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/227,108
; FILING DATE: 03-APR-1994
; CLASSIFICATION: 435
; ATTORNEY/AGENT INFORMATION:
; NAME: Parucci, Allan A.
; REGISTRATION NUMBER: 30,256
; REFERENCE/DOCKET NUMBER: 7620-033
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 212 790-9090
; TELEFAX: 212 869-8864/9741
; TELEX: 66141 PENNIE
; INFORMATION FOR SEQ ID NO: 3:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 379 amino acids
; TYPE: amino acid
; TOPOLOGY: linear
; MOLECULE TYPE: protein
US-08-227-108-3

Query Match 50.8%; Score 1141.5; DB 1; Length 379;
Best Local Similarity 56.5%; Pred. No. 4.4e-112;
Matches 210; Conservative 67; Mismatches 90; Indels 5; Gaps 3;

Qy 37 PTKAVDPEAFNISEIIHQGYPCPEYEVEATEDGYILSVNRIIPRGLVQPKKTSRPPVLL 96
Db 7 PT---NPEVTNISIOMITYMGYPABEYEVVTEDEGYILGIDRIPYGRKNSENIGRRPVAFL 63
Qy 97 QHGLVGGASNWSNLPPNSLGFILADAGFDVWMCNSRGNWSKHKHTLSIDQDEFWAFSY 156
Db 64 QHGLASATNWSNLPPNSLAFILADAGYDVWLCNSRGNTWARNLNLYSPDSVEFWAFSF 123
Qy 157 DEMARFDPVAVINFILQKTEKIIYVYSGTGTMGFIAPSTWPELAQKIMYFALAPIA 216
Db 124 DEMAKYDLPATIDFLLKTKTGDKLHYVGHSGTGTIGFIAPSTWPELAQKIMYFALAPIA 183
Qy 217 TVKHAQSPGTFKLLPDMMLKGLFGKKEFLYQTRFLQ-LVIYLCGVQLDQCSNIMLL 275
Db 184 TVKTEILLNKLMLVPFSLFKLIFGNKIP-YPHHFFDQFLATEVCSRETVDLLCSNALFI 242
Qy 276 LGGFNTNMMNSRASVYAAHTLACTSVQNTILHWSQAVNSGELRAFDPWGSSTKLEKCNQPT 335
Db 243 ICGFDTNMLNSRLDVTYVSHNPAGTSVQNLHWSQAVKSGKQAFDWSGVPQNMHHQS 302
Qy 336 TPVRYRVDMTVPTAMTGGQDMLSNPDEVKMLSEVTNLIYHKNIPDEWAHVDPIWGLDA 395
Db 303 MPYYNTLDMHVPVAVWSGGNDLADPHDVLNLSKLPNLIYHKKIPPPYHNLDFIWMADA 362

Qy 396 PHRMYN E I H L M 407
| : | | | : : |
Db 363 POAVYN E I V S M M 374

RESULT 8

```

US-09-073-674-3
; Sequence 3, Application US/09073674
; Patent No. 5998189
; GENERAL INFORMATION:
; APPLICANT: Blanchard, Claire
; APPLICANT: Benicourt, Claude
; APPLICANT: Junien, Jean-Louis
; TITLE OF INVENTION: Recombinant Dog Gastric Lipase
; NUMBER OF SEQUENCES: 21
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Warner-Lambert Company
; STREET: 2800 Plymouth Road
; CITY: Ann Arbor
; STATE: Michigan
; COUNTRY: U.S.A.
; ZIP: 48105
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patent in Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/09/073,674
; FILING DATE:
; CLASSIFICATION:
; ATTORNEY/AGENT INFORMATION:
; NAME: Crissey, Todd M.
; REGISTRATION NUMBER: 37,807
; REFERENCE/DOCKET NUMBER: 5072-DI-66-TWC
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 734 622-7530
; TELEFAX: 734 622-1553
; TELEX:
; INFORMATION FOR SEQ ID NO: 3:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 379 amino acids
; TYPE: amino acid
; TOPOLOGY: linear
; MOLECULE TYPE: protein
US-09-073-674-3

```

Query Match	50.8%;	Score 1141.5;	DB 2;	Length 379;
Best Local Similarity	56.5%;	Pred. No. 4.4e-112;		
Matches 210;	Conservative 67;	Mismatches 90;	Indels 5;	Gaps 31
QY	37	PTKAVDPEAFNMISIIIOHQYPCPEEYEVATEDGYILSVNRIGLVQPKKGTGSRPVVLL	96	
Db	7	PT---NPEVTNNISOMITWYGPAAEEYEVATEDGYILGIDRIPIYGRKNSENIGRPVAPL	63	
QY	97	QHGLVGGASNTSNI PNNSLGFILADAGFDVVMGNSRGNAWSRKHTLSIDODEFWAIFY	156	
Db	64	QHGLLASATNTSNI PNNSLAFILADAGDYVWLGNRSRGTWARRNLYIYSPDSVEFWAIFY	123	
QY	157	DEMAKFDLPVINFILQKTGQEKIYYGVYSGTITMGFIASFSTWPELAKIKWFEALAPTA	216	
Db	124	DEWAKYDLPATIDFLIKTGTQKLYVGHSGTITIGFIASFSTWPKLAKIKTFYALAPVA	183	
QY	217	TVYHAKSPGTKEFLLPDMWIKGL FOKKEFLYQTRFLRQ-LVIYLCGGVILDOICSNIMLL	275	
Db	184	TVKYTETLLKMLVPSEFLKLIIGNKTF-YPHFFPDQFLATVCSRGTVDLLCSNALPI	242	
QY	276	LGSENTNNMWSRA SVYAHTLAGTSVQNIILHWSAVNSGELAFDWSGSETKNLEKQCP	335	
Db	243	ICGFDTNMLNWSRLDVIYLSHPAGTSVQNVILHWSQAVKSGKQFQFDWGSFVQNMHHYGS	302	
QY	336	TPVRYVRDMTYPTAMWTGGQDMLSNPNEDVKNLLSEVTNLIYFKNIPPEWAHVDFIWLGDA	395	

Db 303 MPPEYNLTDMHYPIAVWNGENDLLADPHDVLILLSKLPNLIYHRKIPPYNHLDFIWANDA 362

QY 396 PRRMYNEIIHLM 407

Db 363 POAVNEIIVSM 374

RESULT 9

```

US-09-348-930A-2
; Sequence 2, Application US/09348930A
; Patent No. 6573431
; GENERAL INFORMATION:
; APPLICANT: Lenee, P
; APPLICANT: Gruber, V
; APPLICANT: Baudino, S
; APPLICANT: Merot, B
; APPLICANT: Benicourt, C
; APPLICANT: Cudrey, C
; TITLE OF INVENTION: Recombinant Predudena,
; FILE REFERENCE: 18433/2012
; CURRENT APPLICATION NUMBER: US/09/348,930A
; CURRENT FILING DATE: 1999-07-02
; PRIOR APPLICATION NUMBER: 08/945,321
; PRIOR FILING DATE: 1998-02-12
; PRIOR APPLICATION NUMBER: PCT/FR96/00606
; PRIOR FILING DATE: 1996-03-19
; NUMBER OF SEQ ID NOS: 37
; SOFTWARE: Patent in version 3.0
; SEQ ID NO 2
; LENGTH: 379
; TYPE: PRT
; ORGANISM: Canis familiaris
US-09-348-930A-2

```

Query Match 50.8%; Score 1141.5; DB 4; Length 379;
Best Local Similarity 56.5%; Pred. No. 4.4e-112;
Matches 210; Conservative 67; Mismatches 90; Indels 5;

Qy	37	PTKAVDPRAFNWISIIHQOYPCPEEYEVATEGYILSVNRIPRGLVPKQKTSRRPVLL	96
Db	7	PT---NPEVTNMISSITMYGPAEEYEVTEGYILGIDRIPYGRKNSENIGRRRVAEL	63
Qy	97	QHGLVCGASNIWISLNPNSLGFILLADAGFVWMCNSRGNAWSRKHTLSIDODEWAPSY	156
Db	64	QHGLASATNMIWISLNPNSLAFILLADAGYDVLGNSRGNWTWARRNIYSPDSVEFWAFSF	123
Qy	157	DEWARFDLPVAVINFILQKTGOBKLYYGYSGQITMGFIAPSTPBLAQKIMYFALAPIA	216
Db	124	DEWAKYDLPATIDFLIKTGGDKLHYVGHSGQITIGFIAPSTNPKLAKIKTYFALAPYA	183
Qy	217	TVYHAKSPGTRFLLPDMMIKGLFKKKEFYQTRFLRQ-LVYILCGQVILDOICSNMILL	275
Db	184	TVKYTETLLNKLMLVPSPLFKLIFGNKIF-YPHHFPDQFLATEVCVSRGTVDLLCSNALFI	242
Qy	276	LGGFTNNWMSRASVYAAHTLACTSVONILHWSQAVNSGELPAFPWGSETKOLEKCNOP	335
Db	243	ICGFDTMLNMSRLDVLVLSHPAGTSYQNVLHWSQAVKSGKFOAFPWGSPVQNMHHYHOS	302
Qy	336	TPVRYVRDMTVPYTAAMTGGQDWLSNPEDVKMLLSEVTNLIYHKNTPEWARVDFIWGLDA	395
Db	303	MPZYNTLDMEVPVAVNNGNDLLADPHVDLLLSKLPNLIYRKLTPPYNHLDFTWAMDA	362
Qy	396	PHRMVNEIHLM	407
Db	363	POAVTNEIYVMM	374

RESULT 20

US-08-227-108-5
; Sequence 5, Application US/08227108
; Patent No. 5807726
; GENERAL INFORMATION;

[illegible]

RESULT 12

US-09-186-489-2
 ; Sequence 2, Application US/09186489
 ; Patent No. 6375947
 ; GENERAL INFORMATION:
 ; APPLICANT: Bolen, Paul L.
 ; APPLICANT: Chak, Paul L.
 ; APPLICANT: Scharf Jr., Lewis G.
 ; TITLE OF INVENTION: Purified Recombinant Kid Pregastric Esterase, and
 ; TITLE OF INVENTION: Processes for its Production and Use
 ; FILE REFERENCE: 5499/3
 ; CURRENT APPLICATION NUMBER: US/09/186,489
 ; CURRENT FILING DATE: 1998-11-05
 ; NUMBER OF SEQ ID NOS: 7
 ; SOFTWARE: Patent in Ver. 2.0
 ; SEQ ID NO 2
 ; LENGTH: 378
 ; TYPE: PRT
 ; ORGANISM: Kid (Goat)
 ; US-09-186-489-2

Query Match 50.2%; Score 1127; DB 4; Length 378;
 Best Local Similarity 55.1%; Pred. No. 1.5e-110;
 Matches 205; Conservative 66; Mismatches 101; Indels 0; Gaps 0;
 QY 40 AVDEAFPMNISEIIHQGYPCSEYEVATEDGYILSVNRIPRGLVQPKTKGSRPVVLLQHG 99
 Db 6 AKNPEASMNVSQMSIFWGYPSEMHKVTADGYILQVYRIPHGKNDANHLGQRPVFLQHG 65
 QY 100 LVGCASWISNLPNNSLGFILADAGFDVWNGSRGNARSKKHTLSIDODEFWAFSYDEM 159
 Db 66 LLASATNWSNLPNNSLGFILADAGFDVWNGSRGNARSKKHTLSIDODEFWAFSYDEM 125
 QY 160 ARFDLPVINFILQKQGOEKIYVYVGSQGTMTGFIATFSTMPPELAQIKMYFALAPATVK 219
 Db 126 AEYDLPSTIDFILKRTQKGLHYVGHSGQTTIGFVAFSTNPTLAEKIEVHALAPATVK 185
 QY 220 HAKSPGTYKFLLLPDMIMKGLFGKKEFLYQTRFRLQVLYLGGVILDDQICSNMILLGGF 279
 Db 186 HTQSLFNKALIPHLFKIIFGNKMFYPHNFQFLGVEVCSRETLDVLCNKALFAITGA 245
 QY 280 NTNNMNSRASVYAAHTLAGTSVQNLHWSQVNSGELRAFDMGSETKLEKCNQPTPV 339
 Db 246 DNKNFNSRLDYYVAHNPAGASVQNLHWRQAIKSKGFOAFDNGASVENLMHYNQTPPI 305
 QY 340 YVRDVTPTAMTGGQDMLSNPDEYKLLSEVTNLIYHKNIPWAHVDPIGLDAPHRM 399
 Db 306 YNLTMNVPTAVNSAGQDLLADPDQVDLLSKLSNLHKEIPNYNHLDFIWMADAPQEV 365
 QY 400 YNEIILHMOOEE 411
 Db 366 YNEIILMAKDK 377

RESULT 13
 US-10-043-665B-2
 ; Sequence 2, Application US/10043665B
 ; Patent No. 6582948
 ; GENERAL INFORMATION:
 ; APPLICANT: Bolen, Paul L.
 ; APPLICANT: Chak, Paul L.
 ; APPLICANT: Scharf Jr., Lewis G.
 ; TITLE OF INVENTION: Recombinant Kid Pregastric Esterase and Methods for Its
 ; TITLE OF INVENTION: Production and Use
 ; FILE REFERENCE: IFF-0009
 ; CURRENT APPLICATION NUMBER: US/10/043,665B
 ; CURRENT FILING DATE: 2002-09-24
 ; PRIOR APPLICATION NUMBER: US 09/186,489
 ; PRIOR FILING DATE: 1998-11-05
 ; NUMBER OF SEQ ID NOS: 7
 ; SOFTWARE: Patent in Ver. 2.0
 ; SEQ ID NO 2
 ; LENGTH: 378
 ; TYPE: PRT

; ORGANISM: Kid (Goat)
 US-10-043-665B-2
 Query Match 50.2%; Score 1127; DB 4; Length 378;
 Best Local Similarity 55.1%; Pred. No. 1.5e-110;
 Matches 205; Conservative 66; Mismatches 101; Indels 0; Gaps 0;
 QY 40 AVDEAFPMNISEIIHQGYPCSEYEVATEDGYILSVNRIPRGLVQPKTKGSRPVVLLQHG 99
 Db 6 AKNPEASMNVSQMSIFWGYPSEMHKVTADGYILQVYRIPHGKNDANHLGQRPVFLQHG 65
 QY 100 LVGCASWISNLPNNSLGFILADAGFDVWNGSRGNARSKKHTLSIDODEFWAFSYDEM 159
 Db 66 LLASATNWSNLPNNSLGFILADAGFDVWNGSRGNARSKKHTLSIDODEFWAFSYDEM 125
 QY 160 ARFDLPVINFILQKQGOEKIYVYVGSQGTMTGFIATFSTMPPELAQIKMYFALAPATVK 219
 Db 126 AEYDLPSTIDFILKRTQKGLHYVGHSGQTTIGFVAFSTNPTLAEKIEVHALAPATVK 185
 QY 220 HAKSPGTYKFLLLPDMIMKGLFGKKEFLYQTRFRLQVLYLGGVILDDQICSNMILLGGF 279
 Db 186 HTQSLFNKALIPHLFKIIFGNKMFYPHNFQFLGVEVCSRETLDVLCNKALFAITGA 245
 QY 280 NTNNMNSRASVYAAHTLAGTSVQNLHWSQVNSGELRAFDMGSETKLEKCNQPTPV 339
 Db 246 DNKNFNSRLDYYVAHNPAGASVQNLHWRQAIKSKGFOAFDNGASVENLMHYNQTPPI 305
 QY 340 YVRDVTPTAMTGGQDMLSNPDEYKLLSEVTNLIYHKNIPWAHVDPIGLDAPHRM 399
 Db 306 YNLTMNVPTAVNSAGQDLLADPDQVDLLSKLSNLHKEIPNYNHLDFIWMADAPQEV 365
 QY 400 YNEIILHMOOEE 411
 Db 366 YNEIILMAKDK 377

RESULT 14
 US-09-348-930A-9
 ; Sequence 9, Application US/09348930A
 ; Patent No. 6573431
 ; GENERAL INFORMATION:
 ; APPLICANT: Lenee, P
 ; APPLICANT: Gruber, V
 ; APPLICANT: Baudino, S
 ; APPLICANT: Merot, B
 ; APPLICANT: Benicourt, C
 ; APPLICANT: Cudrey, C
 ; TITLE OF INVENTION: Recombinant Preducodenal Lipases and Polypeptides Derivatives Prod
 ; TITLE OF INVENTION: Plants, Processes for Obtaining Them and Their Uses
 ; FILE REFERENCE: 18433/2012
 ; CURRENT APPLICATION NUMBER: US/09/348,930A
 ; CURRENT FILING DATE: 1999-07-02
 ; PRIOR APPLICATION NUMBER: 08/945,321
 ; PRIOR FILING DATE: 1998-02-12
 ; PRIOR APPLICATION NUMBER: PCT/FR96/00606
 ; PRIOR FILING DATE: 1996-03-19
 ; NUMBER OF SEQ ID NOS: 37
 ; SOFTWARE: Patent in version 3.0
 ; SEQ ID NO 9
 ; LENGTH: 398
 ; TYPE: PRT
 ; ORGANISM: Homo sapiens
 ; US-09-348-930A-9

Query Match 49.8%; Score 1119; DB 4; Length 398;
 Best Local Similarity 52.2%; Pred. No. 1.1e-109;
 Matches 210; Conservative 73; Mismatches 107; Indels 12; Gaps 2;
 QY 17 MWLLILVAYMFQNVNSVEMPTKAV-----DPEAFMNISETIHOQGYPCREYEVATED 69
 Db 1 MWLLITWASLI-----SVLGTTHGLFGKHLPGSFVTMNIQSMTIWTGTFNNEIEYVTTED 55
 QY 70 GYILSVNRIPRGLVQPKTKGSRPVVLLQHGVLGASNWSINLPNNSLGFILADAGFDVW 129

Db 56 GYILEVNRIPYKKNKSGNTQRPVFLQHGLLASATNWSLNPNSLAFILADAGYDVL 115
 QY 130 GNSRGNAWSKHKHTLSIDQDEWAFSDENAREDLPAVINFILOKTOGQEKIYVGYSGQT 189
 Db 116 GNSRGNTWARNLYSPDSVEFWAFSDENAREDLPAVINFILOKTOGQEKIYVGYSGQT 175
 QY 190 TMGFIAFTMPLELAQKIMYFALAPIATVHAQSPGTFKLLPDMIMKGLFGKGEFLYQT 249
 Db 176 TIGFIAFTNPSLAKRIKTFYALAPVATVYTKSLINKLRFVPSLKFIFGDKIPYPHN 235
 QY 250 RELRQVILYLCGQVILDOICSNIMLLGGENTNNMNSRASVYAAHTLAGTSVONILHWS 309
 Db 236 FFOQFLATEVCSREMLNLLCSNLFICGPDSCNFTSLDVYLSHNPAGTSVQNNFHT 295
 QY 310 QAVNSGELRAFDRGSETKNEKCNQPTFVRYRVRDMTVPMTAMTGGQDMLSNPDEYKMLL 369
 Db 296 QAVKSGKGFQAVDWSVPQNRMHYDQSQPPYVYVAMNVPIAVNNGKDLADPDQVGLLL 355
 QY 370 SEVTNLIYHKNIPENAHVDPIWGLDAPHRMNIEIHLMOQEE 411
 Db 356 PKLPNLIYHKEIPFYNHLDPIWADAPQEVYNDIVSMISEDK 397

RESULT 15

US-08-227-108-16
 ; Sequence 16 Application US/08227108
 ; Patent No. 5807726
 ; GENERAL INFORMATION:
 ; APPLICANT: Blanchard, Claire
 ; APPLICANT: Benicourt, Claude
 ; APPLICANT: Junien, Jean-Louis
 ; TITLE OF INVENTION: Recombinant Dog Gastric Lipase
 ; NUMBER OF SEQUENCES: 21
 ; CORRESPONDENCE ADDRESSES:
 ; ADDRESSEE: Pennie & Edmonds
 ; STREET: 1155 Avenue of the Americas
 ; CITY: New York
 ; STATE: New York
 ; COUNTRY: U.S.A.
 ; ZIP: 10036
 ; COMPUTER READABLE FORM:
 ; MEDIUM TYPE: Floppy disk
 ; COMPUTER: IBM PC compatible
 ; OPERATING SYSTEM: PC-DOS/MS-DOS
 ; SOFTWARE: Patent In Release #1.0, Version #1.25
 ; CURRENT APPLICATION DATA:
 ; APPLICATION NUMBER: US/08/227,108
 ; FILING DATE: 03-APR-1994
 ; CLASSIFICATION: 435
 ; ATTORNEY/AGENT INFORMATION:
 ; NAME: Panucci, Allan A.
 ; REGISTRATION NUMBER: 30,256
 ; REFERENCE/DOCKET NUMBER: 7620-033
 ; TELECOMMUNICATION INFORMATION:
 ; TELEPHONE: 212 790-9090
 ; TELEFAX: 212 869-8864/9741
 ; TELEX: 66141 PENNIE
 ; INFORMATION FOR SEQ ID NO: 16:
 ; SEQUENCE CHARACTERISTICS:
 ; LENGTH: 380 amino acids
 ; TYPE: amino acid
 ; STRANDEDNESS: single
 ; TOPOLOGY: linear
 ; MOLECULE TYPE: protein
 ; US-08-227-108-16

Query Match 48.9% Score 1098.5; DB 1; Length 380;
 Best Local Similarity 54.6%; Pred. No. 1.5e-107;
 Matches 202; Conservative 69; Mismatches 98; Indels 1; Gaps 1;
 QY 43 PEAPMNISETIQHGYPCBEYEVATEDGYILSVNRPGLVQPKTKGSRPVLLOHGLVG 102

Db 10 PEVTMNISQMITYMGVYNEEYVVTEDGYILEVNRIPYKKNKSGNTQRPVFLQHGLLA 69
 QY 103 GASNWSLNPNSLFGFTILADAGDVWNGSRGNWSKHKHTLSIDQDEFW-AFSYDEMAR 161
 Db 70 SATNWSLNPNSLAFILADAGDVWNGSRGNWTWARNLYYSPDSVEFWAFSDENAK 129
 QY 162 FDLPAVINFILOKTOGQEKIYVGYSGQTTFATFSTMPLELAQKIMYFALAPIATVHA 221
 Db 130 YDLPATIDFIVKKTGQKQLHYVGHSGQTTFATFSTNPSLAKRIKTFYALAPVATVKYT 189
 QY 222 KSPGTFKLLPDMIMKGLFGKGEFLYQTRFLRQVILYLCGQVILDOICSNIMLLGGFNT 281
 Db 190 KSLINKLRFVPSLKFIFGDKIPYPHNFPDQFLATEVCSREMLNLLCSNLFICGDS 249
 QY 282 NNMNSRASVYAAHTLAGTSVONILHWSQAVNSGELRAFDRGSETKNEKCNQPTFVRYR 341
 Db 250 KNFTSLRLDVYLSHNPAGTSVQNNFHTQAVKSGKGFQAVDWSVPQNRMHYDQSQPPYV 309
 QY 342 VRDMTVPMTAMTGGQDMLSNPDEYKMLLSEVTNLIYHKNIPENAHVDPIWGLDAPHRMYN 401
 Db 310 VTAMNVPIAVNNGKDLADPDQVGLLLPKLPNLIYHKEIPFYNHLDPIWADAPQEVYN 369
 QY 402 EIIHLMQEE 411
 Db 370 DIVSMISEDK 397

Search completed: February 19, 2004, 14:23:46
 Job time: 16.5683 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2004 Compugen Ltd.

OM protein - protein search, using sw model

Run on: February 19, 2004, 14:16:19 ; Search time 9.11439 Seconds
(without alignments)
2012.248 Million cell updates/sec

Title: US-10-042-431-49

Perfect score: 2076

Sequence: 1 VHMPTKAYDPAFMNISEII.....IHLMQBETNLISQRCRAVL 390

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 127863 seqs, 47026705 residues

Total number of hits satisfying chosen parameters: 127863

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : SwissProt_41.*

Pred. No. is the number of results predicted by chance to have a
score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	1289	62.1	399	1	LICH HUMAN
2	1220.5	58.8	397	1	LICH RAT
3	1194.5	57.5	397	1	LICH MOUSE
4	1161	55.9	395	1	LIPG RAT
5	1141.5	55.0	398	1	LIPG CANFA
6	1125	54.2	397	1	LIPG BOVIN
7	1109	53.4	398	1	LIPG HUMAN
8	663	31.9	394	1	LIP3 DROME
9	546	26.3	439	1	LIP1 DROME
10	422.5	20.4	548	1	TGL1 YEAST
11	111.5	5.4	987	1	YD94 METUA
12	105	5.1	430	1	ORC2 CABEL
13	98.5	4.7	436	1	PAPA CAVPO
14	96.5	4.6	287	1	Y193 HAEIN
15	93.5	4.5	278	1	PRXC STRAU
16	93	4.5	569	1	PYRD PLAPA
17	92	4.4	748	1	PTIP ECOLI
18	92	4.4	748	1	PTIP SALTY
19	91	4.4	370	1	P2C2 SCHPO
20	90	4.3	277	1	BPA2 STRAU
21	89	4.3	330	1	EXOA RHINE
22	88.5	4.3	185	1	Y2H8 BACSU
23	88	4.2	1259	1	L1N1 HUMAN
24	87	4.2	473	1	SCRB LACIA
25	86.5	4.2	382	1	YVHA HAEIN
26	86.5	4.2	685	1	INVA SALTY
27	86.5	4.2	869	1	CPAC ECOLI
28	86.5	4.2	1122	1	Y383 YEAST
29	86	4.1	422	1	Y383 SHEON
30	86	4.1	1866	1	VGNB CPMV
31	85.5	4.1	428	1	Y563 SCHPO
32	85	4.1	292	1	FAEB NEUCR
33	85	4.1	429	1	ARB2_ECOLI

RESULT 1

ID	LICH HUMAN	STANDARD;	PRT;	399 AA.
AC	P38571; Q16529; Q96EJO;			
DT	01-OCT-1994 (Rel. 30, Created)			
DT	01-OCT-1994 (Rel. 30, Last sequence update)			
DT	15-SEP-2003 (Rel. 42, Last annotation update)			
DE	Lysosomal acid lipase/cholesterol ester hydrolase precursor			
DE	(EC 3.1.1.13) (LAL) (Acid cholesterol ester hydrolase) (Sterol			
DE	esterase) (Lipase A) (Cholesteryl esterase).			
GN	LIPA.			
OS	Homo sapiens (Human).			
OC	Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;			
OC	Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.			
OX	NCBI_TaxID=9606;			
RN	[1]			
RP	SEQUENCE FROM N.A., AND SEQUENCE OF 196-212; 277-297 AND 305-315.			
RX	MEDLINE=92042192; PubMed=1718995;			
RA	Anderson R.A.; Sando G.N.;			
RT	"Cloning and expression of cDNA encoding human lysosomal acid			
RT	lipase/cholesterol ester hydrolase. Similarities to gastric and			
RT	lingual lipases.";			
RL	J. Biol. Chem. 266:22479-22484 (1991).			
RN	[2]			
RP	SEQUENCE FROM N.A.			
RX	TISSUE=Liver;			
RX	MEDLINE=94155897; PubMed=8112342;			
RA	Ameis D., Merkel M., Eckerskorn C., Greten H.;			
RT	"Purification, characterization and molecular cloning of human			
RT	hepatic lysosomal acid lipase.";			
RL	Eur. J. Biochem. 219:905-914 (1994).			
RN	[3]			
RP	SEQUENCE FROM N.A.			
RX	TISSUE=Liver;			
RX	MEDLINE=96363957; PubMed=8725147;			
RA	Du H., Witte D.P., Grabowski G.A.;			
RT	"Tissue and cellular specific expression of murine lysosomal acid			
RT	lipase mRNA and protein.";			
RL	J. Lipid Res. 37:937-949 (1996).			
RN	[4]			
RP	SEQUENCE FROM N.A.			
RX	TISSUE=Placenta;			
RX	MEDLINE=22388257; PubMed=12477932;			
RA	Klausner R.D., Collins F.S., Wagner L., Shenmen C.M., Schuler G.D.,			
RA	Staudenmann R.D., Buetow K.H., Schaefer C.F., Bhat N.K.,			
RA	Altschul S.F., Zeeberg B., Buetow K.H., Moore T., Max S.I., Wang J., Hsieh P.,			
RA	Hopkins R.F., Jordan H., Moore T., Max S.I., Wang J., Hsieh P.,			
RA	Diatchenko L., Marusina K., Farmer A.A., Rubin G.M., Hong L.,			
RA	Stapleton M., Soares M.B., Bonaldo M.F., Casavant T.L., Schetz T.B.,			
RA	Brownstein M.J., Usdin T.B., Toshiyuki S., Carninci P., Prange C.,			
RA	Raha S.S., Loquellano N.A., Peters G.J., Abramson R.D., Mullaly S.J.,			
RA	Bosak S.A., McEwan P.J., McKernan K.J., Malek J.A., Gunaratne P.H.,			
RA	Richards S., Worley K.C., Hale S., Garcia A.M., Gay L.J., Hulyk S.W.,			
RA	Villalon D.K., Muzny D.M., Sodergren E.J., Lu X., Gibbs R.A.,			
RA	Fahey J., Helton E., Kettner M., Madan A., Rodriguez S., Sanchez A.,			
RA	Whiting M., Madan A., Young A.C., Shevchenko Y., Bouffard G.G.,			

ALIGNMENTS

Q89ak4 buchnera ap
Q9uuJ9 schizosacch
P97313 mus musculus
P38295 saccharomyc
P36430 bacillus su
P52408 prunus pers
Q9QYF9 mus musculus
P55882 salmonella
Q31158 pseudomonas
Q9KVB1 vibrio chol
P04635 staphylococ
P16115 thermotoga

34 85 4.1 429 1 BIOA BUCBP
35 85 4.1 521 1 ACHI SCHPO
36 85 4.1 4128 1 PRXD MOUSE
37 84 4.0 451 1 YB27 YEAST
38 84 4.0 804 1 SYL BACSU
39 83.5 4.0 350 1 E13B PRUPE
40 83.5 4.0 375 1 NDR3 MOUSE
41 83 4.0 266 1 THID SALTY
42 83 4.0 273 1 PRXC PSEPL
43 83 4.0 429 1 PUR2 VIBCH
44 83 4.0 641 1 LIP STAFY
45 82.5 4.0 319 1 LDH THEMA

RA Blakesley R.W., Touchman J.W., Green E.D., Dickson M.C.,
 RA Rodriguez A.C., Grimwood J., Schmitz J., Myers R.M.,
 RA Butterfield Y.S.N., Krzywinski M.I., Skalska U., Smalilus D.E.,
 RA Schnerch A., Schein J.E., Jones S.U.M., Marra M.A.;
 RT "Generation and initial analysis of more than 15,000 full-length
 RT human and mouse cDNA sequences.";
 RL Proc. Natl. Acad. Sci. U.S.A. 99:16899-16903(2002).
 RN [5]
 RP VARIANT CESD/WD PRO-200.
 RX MEDLINE-94195814; PubMed-8146180;
 RA Anderson R.A., Byrum R.S., Coates P.M., Sando G.N.;
 RT "Mutations at the lysosomal acid cholesteryl ester hydrolase gene
 RT locus in Wolman disease.";
 RL Proc. Natl. Acad. Sci. U.S.A. 91:2718-2722(1994).
 RN [6]
 RP VARIANTS CESD ARG-129 and PRO-129.
 RX MEDLINE-98235576; PubMed-9633819;
 RA Ries S., Buechler C., Schindler G., Alanidis C., Aneis D., Gasche C.,
 RA Jung N., Schambach A., Fehrer P., Vanier M.T., Belli D.C.,
 RA Greten H., Schmitz G.;
 RT "Different missense mutations in histidine-108 of lysosomal acid
 RT lipase cause cholesteryl ester storage disease in unrelated compound
 RT heterozygous and hemizygous individuals.";
 RL Hum. Mutat. 12:44-51(1998).
 CC -1- FUNCTION: CRUCIAL FOR THE INTRACELLULAR HYDROLYSIS OF CHOLESTERYL
 CC ESTERS AND TRIGLYCERIDES THAT HAVE BEEN INTERNALIZED VIA RECEPTOR-
 CC MEDIATED ENDOCYTOSIS OF LIPOPROTEIN PARTICLES. IMPORTANT IN
 CC MEDIATING THE EFFECT OF LDL (LOW DENSITY LIPOPROTEIN) UPTAKE ON
 CC SUPPRESSION OF HYDROXYMETHYLGLUTARYL-COA REDUCTASE AND ACTIVATION
 CC OF ENDOGENOUS CELLULAR CHOLESTERYL ESTER FORMATION.
 CC -1- CATALYTIC ACTIVITY: A sterol ester + H(2)O = a sterol + a fatty
 CC acid
 CC -1- SUBCELLULAR LOCATION: Lysosomal.
 CC -1- DISEASE: DEFECTS IN LIPA ARE THE CAUSE OF THE SEVERE INFANTILE-
 CC ONSET WOLMAN DISEASE (WD) AND THE Milder LATE-ONSET CHOLESTERYL
 CC ESTER STORAGE DISEASE (CESD).
 CC -1- SIMILARITY: BELONGS TO THE ABHYDROLASE SUPERFAMILY. LIPASE FAMILY.
 CC
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (see <http://www.isb-sib.ch/announce/>
 CC or send an email to license@sib-sib.ch).
 CC
 DR EMBL; M74775; AAA59519.1; -;
 DR EMBL; U04285; AAB60327.1; -;
 DR EMBL; U04286; AAB60327.1; JOINED.
 DR EMBL; U04287; AAB60327.1; JOINED.
 DR EMBL; U04288; AAB60327.1; JOINED.
 DR EMBL; U04290; AAB60327.1; JOINED.
 DR EMBL; U04291; AAB60327.1; JOINED.
 DR EMBL; U04292; AAB60327.1; JOINED.
 DR EMBL; U04293; AAB60327.1; JOINED.
 DR EMBL; X76488; CAA54026.1; -;
 DR EMBL; X31690; CAA83495.1; -;
 DR EMBL; U08464; AAB60328.1; -;
 DR EMBL; BC012287; AAH12287.1; -;
 DR PIR; S41408; S41408.
 DR Genew; HGNC:6617; LIPA.
 DR MIM; 278000; -;
 DR GO; GO:0005764; C:lysosome; TAS.
 DR GO; GO:0006487; P:N-linked glycosylation; TAS.
 DR InterPro; IPR000073; A/b hydrolase.
 DR InterPro; IPR000734; Lipase.
 DR InterPro; IPR000379; Ser esterase site.
 DR Pfam; PF04083; abhydrolase_1.
 DR Pfam; PF00561; abhydrolase; 1.
 DR PROSITE; PS00120; LIPASE_SER; 1.
 KW Hydrolase; lipid degradation; Glycoprotein; Signal; Lysosome;
 KW Disease mutation; Polymorphism.
 FT SIGNAL 1 21
 FT POTENTIAL.

FT CHAIN 22 399
 FT ACT_SITE 174 174
 FT ACT_SITE 374 374
 FT CARBOHYD 36 36
 FT CARBOHYD 72 72
 FT CARBOHYD 101 101
 FT CARBOHYD 161 161
 FT CARBOHYD 273 273
 FT CARBOHYD 321 321
 FT VARIANT 16 16
 FT VARIANT 129 129
 FT VARIANT 129 129
 FT VARIANT 200 200
 FT CONFLICT 23 23
 FT CONFLICT 29 29
 SQ SEQUENCE 399 AA; 45415 MW; 55F27391306B09A CRC64;
 Query Match 62.1%; Score 1289; DB 1; Length 399;
 Best Local Similarity 63.5%; Pred. No. 7.6e-107;
 Matches 235; Conservative 55; Mismatches 80; Indels 0; Gaps 0;
 QY 7 AVDEAFNISEIIHQGYPCCEVEVATEDGYLLSVNRIIPRGLVQPKTGSRRVYLLQHG 66
 DB 28 AVDEPTNNVSEIISYWGFPSEYLVETEDGYLLSVNRIIPRGLVQPKTGSRRVYLLQHG 87
 QY 67 LVGASWISNLPNNSLGFILADAGFDVWNGSNGRANWSKHKHTLSIDQDEFWAFSYDEM 126
 DB 88 LLADSSNWTNLSNLSLGFILADAGFDVWNGSNGRANWSKHKHTLSIDQDEFWAFSYDEM 147
 QY 127 ARFDLPVINFILQKTOEKIYVGVSGQTMGFIAPSTMPPELAQIKTIFALAPIATVK 186
 DB 148 AKYDLPASINFILNKTKQEQVYVGHSGQTTIGFIAPSQIPELAKIKMFALGPVASVA 207
 QY 187 HAKSPGTFKLLPDMMLTKGLFKGKELVQTRFLQVLYLGVGVILDCVSNMLLGGF 246
 DB 208 FCTSPMAKGLRPLDLKDFGDKFELPQSAFLKWLGVHVTHTVILKELGNCNLCFLCGP 267
 QY 247 NTNNMNSRASVYAAHTLAGTSVQNLHWSQAVNSGELRAFDMGSETKNEKCNQPTFVR 306
 DB 268 NERNLNSRVDVYTHSPAGTSVQNLHWSQAVKFKQFAFDWGSAAKNTFYHNSQSYPT 327
 QY 307 YRVDMVTPTAMWTGGODWLSNPEDVZKLLSEVNTLYHKNIPEWAHVDEINGLDAPHRM 366
 DB 328 YNVKMLVPTAVWSGGHDLADVYDVNILLTQITNLVFNHESIPWEHLDFINGLDAPWRL 387
 QY 367 YNEIHLMOQ 376
 DB 388 YNKIINLWRK 397
 RESULT 2
 LICH RAT
 ID LICH RAT STANDARD; PRT; 397 AA.
 AC Q64194;
 DT 28-FEB-2003 (Rel. 41, Created)
 DT 28-FEB-2003 (Rel. 41, Last sequence update)
 DT 28-FEB-2003 (Rel. 41, Last annotation update)
 DE Lysosomal acid lipase/cholesteryl ester hydrolase precursor
 DE (EC 3.1.1.13) (LAL) (Acid cholesteryl ester hydrolase) (Sterol
 DE esterase) (Lipase A) (Cholesteryl esterase).
 GN LIPA OR LAL.
 OS Rattus norvegicus (Rat).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Rattus.
 ON NCBI_TaxID=10116;
 RX [1]
 RP SEQUENCE FROM N.A.
 RC TISSUE=Liver;

LYSOSOMAL ACID LIPASE/CHOLESTERYL ESTER
 HYDROLASE.
 CHARGE RELAY SYSTEM (BY SIMILARITY).
 CHARGE RELAY SYSTEM (BY SIMILARITY).
 N-LINKED (GLCNAC. .) (POTENTIAL).
 N-LINKED (GLCNAC. .) (POTENTIAL).
 N-LINKED (GLCNAC. .) (POTENTIAL).
 N-LINKED (GLCNAC. .) (POTENTIAL).
 N-LINKED (GLCNAC. .) (POTENTIAL).
 N-LINKED (GLCNAC. .) (POTENTIAL).
 P -> T.
 /FTId=VAR_004247.
 H -> P (IN CESD).
 /FTId=VAR_004248.
 H -> R (IN CESD).
 /FTId=VAR_004249.
 L -> P (IN CESD AND WD).
 /FTId=VAR_004250.
 G -> R (IN REF. 3).
 V -> L (IN REF. 4).
 55F27391306B09A CRC64;
 45415 MW;
 62.1%; Score 1289; DB 1; Length 399;
 63.5%; Pred. No. 7.6e-107;
 235; Conservative 55; Mismatches 80; Indels 0; Gaps 0;
 7 AVDEAFNISEIIHQGYPCCEVEVATEDGYLLSVNRIIPRGLVQPKTGSRRVYLLQHG 66
 28 AVDEPTNNVSEIISYWGFPSEYLVETEDGYLLSVNRIIPRGLVQPKTGSRRVYLLQHG 87
 67 LVGASWISNLPNNSLGFILADAGFDVWNGSNGRANWSKHKHTLSIDQDEFWAFSYDEM 126
 88 LLADSSNWTNLSNLSLGFILADAGFDVWNGSNGRANWSKHKHTLSIDQDEFWAFSYDEM 147
 127 ARFDLPVINFILQKTOEKIYVGVSGQTMGFIAPSTMPPELAQIKTIFALAPIATVK 186
 148 AKYDLPASINFILNKTKQEQVYVGHSGQTTIGFIAPSQIPELAKIKMFALGPVASVA 207
 187 HAKSPGTFKLLPDMMLTKGLFKGKELVQTRFLQVLYLGVGVILDCVSNMLLGGF 246
 208 FCTSPMAKGLRPLDLKDFGDKFELPQSAFLKWLGVHVTHTVILKELGNCNLCFLCGP 267
 247 NTNNMNSRASVYAAHTLAGTSVQNLHWSQAVNSGELRAFDMGSETKNEKCNQPTFVR 306
 268 NERNLNSRVDVYTHSPAGTSVQNLHWSQAVKFKQFAFDWGSAAKNTFYHNSQSYPT 327
 307 YRVDMVTPTAMWTGGODWLSNPEDVZKLLSEVNTLYHKNIPEWAHVDEINGLDAPHRM 366
 328 YNVKMLVPTAVWSGGHDLADVYDVNILLTQITNLVFNHESIPWEHLDFINGLDAPWRL 387
 367 YNEIHLMOQ 376
 388 YNKIINLWRK 397

RX MEDLINE=96129534; PubMed=8576647;
 RA Nakagawa H., Matsubara S., Kuriyama M., Yoshidome H., Fujiyama J.,
 RA Yoshida H., Oseme M.;
 RT "Cloning of rat lysosomal acid lipase cDNA and identification of the
 RT mutation in the rat model of Wolman's disease.";
 RL J. Lipid Res. 36:2212-2218 (1995).
 CC -!- FUNCTION: CRUCIAL FOR THE INTRACELLULAR HYDROLYSIS OF CHOLESTERYL
 CC ESTERS AND TRIGLYCERIDES THAT HAVE BEEN INTERNALIZED VIA RECEPTOR-
 CC MEDIATED ENDOCYTOSIS OF LIPOPROTEIN PARTICLES. IMPORTANT IN
 CC MEDIATING THE EFFECT OF LDL (LOW DENSITY LIPOPROTEIN) UPTAKE ON
 CC SUPPRESSION OF HYDROXYMETHYLGLUTARYL-COA REDUCTASE AND ACTIVATION
 CC OF ENDOGENOUS CELLULAR CHOLESTERYL ESTER FORMATION (BY
 CC similarity).
 CC -!- CATALYTIC ACTIVITY: A sterol ester + H(2)O = a sterol + a fatty
 CC acid.
 CC -!- SUBCELLULAR LOCATION: Lysosome.
 CC -!- SIMILARITY: BELONGS TO THE ABHYDROLASE SUPERFAMILY. LIPASE FAMILY.
 CC
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@isb-sib.ch).
 CC
 CC EMBL; S81497; AAB36043.2; -;
 CC InterPro; IPR000073; A/b hydrolase.
 CC InterPro; IPR000734; Lipase.
 CC InterPro; IPR000379; Ser esters site.
 CC Pfam; PF04083; abhydro lipase; 1.
 CC Pfam; PF00561; abhydro lipase; 1.
 CC PROSITE; PS00120; LIPASE SER; 1.
 CC PROSITE; PS00120; LIPASE SER; 1.
 CC Hydrolase; Lipid degradation; Signal; Lysosome.
 CC SIGNAL
 CC CHAIN 1 19
 CC LYSOSOMAL ACID LIPASE/CHOLESTERYL ESTER
 CC HYDROLASE.
 CC ACT SITE 172 172
 CC ACT SITE 372 372
 CC CARBOHYD 34 34
 CC CARBOHYD 99 99
 CC CARBOHYD 159 159
 CC CARBOHYD 271 271
 CC CARBOHYD 319 319
 CC SEQUENCE 397 AA; 45186 MW; 97A38595A0523947 CRC64;
 SQ
 Query Match
 Best Local Similarity 58.8%; Score 1220.5; DB 1; Length 397;
 Matches 220; Conservative 67; Mismatches 84; Indels 7; Gaps 3;
 QY 4 FT---KAYDPEAFMNIISIIHQGYPCBEYEVATEDGYLSVNRIPRGLVQPKTGRPV 60
 DB 20 PGTITSAVDPEANMNVTEIIMHWGP--EHSVQICDGYLGVHRIPIHGRKNQPKPKPV 77
 QY 61 VLLQ--HGLVGASNIENLNSLGFILADAGFDVWNGSRGNWAKRKHKLTSIDQDEF 118
 DB 78 VYLRHGFGLADSNVWNIIDNSLGFILADAGFDVWNGSRGNWAKRKHKLTSVSDQY 137
 QY 119 WAFSDENARFDPVAVIFILQKTOEKIYVGVSGQTMGFIATFTWPELAQIKMYEA 178
 DB 138 WAFSDENAKYDLPASINYLKNGQEQQLNVGHSQGTIGFIATFQSPHLLAKVKNFFA 197
 QY 179 LAPATVHAKSPGTFKLLLDPMKIGLFGKKEFLYQTRFRLQRIYVLCGQVILDQICSN 238
 DB 198 LAPVLSLNFASGPMVKGLGELPDLLEDFGQKFLQPSAMVKWLSITCTHVMKELCAN 257
 QY 239 IMLLGGFNTNNMNSRSVVAHAHTLACTSVQNTILHWSQVNSGELRAPDWSGSEKLEK 298
 DB 258 IFFLCGNEKLNLSRVVDVTHCPAGTSVQNVHWTQVVKYHKLQAFDNGSSDKRNFH 317
 QY 299 CNOQTFVRYVRDMVTPTAMVTGGQDLNPNEDVKMLLSVNTLIYHKNIPEWAHVDFIW 358
 DB 318 YNQSYPPLYSIKDMQLPTALWGGKDLADTSDINILLTEIPTLVYHKNIPEWDHDFIW 377

QY 359 GLDAPHRMYNEIHLHMQQ 376
 DB 378 GLDAPRWLYNEVSVLMKK 395

RESULT 3

LICH MOUSE
 ID LICH MOUSE STANDARD; PRT; 397 AA.
 AC QZ0M5;
 DT 28-FEB-2003 (Rel. 41, Created)
 DT 28-FEB-2003 (Rel. 41, Last sequence update)
 DT 28-FEB-2003 (Rel. 41, Last annotation update)
 DE Lysosomal acid lipase/cholesteryl ester hydrolase precursor
 DE (EC 3.1.1.13) (LAL) (Acid cholesteryl ester hydrolase) (Sterol
 DE esterase) (Lipase A) (Cholesteryl esterase).
 GN LIPA OR LIPI.
 OS Mus musculus (Mouse).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
 OX NCBI_TaxID=10090;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN=C57BL/6 X CBA; TISSUE=Liver;
 RX MEDLINE=96363957; PubMed=8725147;
 RA Du H., Witte D.P., Grabowski G.A.;
 RT "Tissue and cellular specific expression of murine lysosomal acid
 RT lipase mRNA and protein.";
 RL J. Lipid Res. 37:937-949 (1996).
 CC -!- FUNCTION: CRUCIAL FOR THE INTRACELLULAR HYDROLYSIS OF CHOLESTERYL
 CC ESTERS AND TRIGLYCERIDES THAT HAVE BEEN INTERNALIZED VIA RECEPTOR-
 CC MEDIATED ENDOCYTOSIS OF LIPOPROTEIN PARTICLES. IMPORTANT IN
 CC MEDIATING THE EFFECT OF LDL (LOW DENSITY LIPOPROTEIN) UPTAKE ON
 CC SUPPRESSION OF HYDROXYMETHYLGLUTARYL-COA REDUCTASE AND ACTIVATION
 CC OF ENDOGENOUS CELLULAR CHOLESTERYL ESTER FORMATION (BY
 CC similarity).
 CC -!- CATALYTIC ACTIVITY: A sterol ester + H(2)O = a sterol + a fatty
 CC acid.
 CC -!- SUBCELLULAR LOCATION: Lysosome.
 CC -!- TISSUE SPECIFICITY: Expressed at low levels in most tissues. High
 CC level expression is found in hepatocytes and splenic and thymic
 CC cells. Very high level expression is observed in cells of the
 CC small intestinal villi, the zona fasciculata and reticularis of
 CC the adrenal cortex, pancreatic acini, and renal tubular
 CC epithelium.
 CC -!- SIMILARITY: BELONGS TO THE ABHYDROLASE SUPERFAMILY. LIPASE FAMILY.
 CC
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@isb-sib.ch).
 CC
 CC EMBL; Z31689; CAA83494.1; -;
 CC MGD; MGI:96789; Lip1.
 CC InterPro; IPR000073; A/b hydrolase.
 CC InterPro; IPR000734; Lipase.
 CC InterPro; IPR000379; Ser esters site.
 CC Pfam; PF04083; abhydro lipase; 1.
 CC Pfam; PF00561; abhydro lipase; 1.
 CC PROSITE; PS00120; LIPASE SER; 1.
 CC PROSITE; PS00120; LIPASE SER; 1.
 CC Hydrolase; Lipid degradation; Signal; Lysosome.
 CC SIGNAL 1 19
 CC LYSOSOMAL ACID LIPASE/CHOLESTERYL ESTER
 CC HYDROLASE.
 CC ACT SITE 172 172
 CC ACT SITE 372 372
 CC CARBOHYD 34 34
 CC CARBOHYD 99 99
 CC CARBOHYD 159 159
 CC CARBOHYD 271 271

FT CARBOHYD 319 319 N-LINKED (GLCNAC. . .) (POTENTIAL).
 SQ SEQUENCE 397 AA; 45551 MW; F896C39E1CCFA91P CRC64;
 Query Match 57.5%; Score 1194.5; DB 1; Length 397;
 Best Local Similarity 57.3%; Pred. No. 1.9e-98;
 Matches 216; Conservative 66; Mismatches 92; Indels 3; Gaps 1;
 QY 3 MPT---KAYDPAFNMISIIIOHGYPCCEYEYVATEDGYILSVNRIPLRGVLPKKTGSRP 59
 DB 19 VPTGTVSADVPVNMVNTIIMRWGYPGEHSLVTCGDGYILSIHRIPLRGVLPKKTGSRP 78
 QY 60 VLLQHLGVGASNTISNIPNNLSGLPILADAGFVMMGNSRGNWSRKHKLISIDQDFW 119
 DB 79 VVYLQHLGLADSNWNTINDSSGLPILADAGFVMMGNSRGNWSRKHKLISVQDFW 138
 QY 120 APSYDMDARFDPVAVINFIQKQEKIYVGVYSGQTTMGTFAPSMPELAQIKMYFAL 179
 DB 139 APSYDMDARFDPVAVINFIQKQEKIYVGVYSGQTTMGTFAPSMPELAQIKMYFAL 198
 QY 180 APIATVKHAKSPOTKFLLLPDMIMKGLFGKGFYQTRFRLQVLYLQGVILQICSN 239
 DB 199 APVLSNFASGFLQLGLPLPDLKDMFGKQFLPOSAMKWLSHVCTHVMKELCANV 258
 QY 240 MLLGGFNTNMMSRASVYAAHTLACTSVQNLHWSQAVNSGELRAPDWSGSETKLEKC 299
 DB 259 FFLCGFNEKLMNSRDVITTHCPAELLVQNLHWSQAVNSGELRAPDWSGSETKLEKC 318
 QY 300 NQPTPVRYRVMTVPTAMWTGQDNLSPEDVLMSEVTLNLYHKNIPWAHVDFIWG 359
 DB 319 NQSPFPYSYKMKRLPTALWGGDRDLNDITILLTQIKLVYHKNIPWHDLDIWI 378
 QY 360 LDAPHMYNEIHLMOQ 376
 DB 379 LDAPWKLYDEIISLMKK 395
 RESULT 4
 LIPG RAT
 ID LIPG RAT STANDARD; PRT; 395 AA.
 AC P04634;
 DT 13-AUG-1987 (Rel. 05, Created)
 DT 13-AUG-1987 (Rel. 05, Last sequence update)
 DT 28-FEB-2003 (Rel. 41, Last annotation update)
 DE Triacylglycerol lipase, lingual precursor (EC 3.1.1.3) (Lingual lipase).
 DE Rattus norvegicus (Rat).
 OC Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;
 OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Rattus.
 OX NCBI_TaxID=10116;
 RN [1]
 RP SEQUENCE FROM N.A.
 RX MEDLINE=85215587; Pubmed=3839077;
 RA Doeherty A.J.P., Bodmer M.W., Angal S., Verger R., Riviere C.,
 RA Lowe P.A., Lyons A., Entage J.S., Harris T.J.R.;
 RT "Molecular cloning and nucleotide sequence of rat lingual lipase
 RT cDNA."
 RL Nucleic Acids Res. 13:1891-1903(1985).
 CC -1- CATALYTIC ACTIVITY: Triacylglycerol + H(2)O = diacylglycerol + a
 CC fatty acid anion.
 CC -1- MISCELLANEOUS: LINGUAL LIPASE IS SECRETED BY THE SEROUS (VON
 CC BERNER'S) GLANDS AT THE BACK OF THE RAT TONGUE.
 CC -1- SIMILARITY: BELONGS TO THE ABHYDROLASE SUPERFAMILY. LIPASE FAMILY.
 CC -----
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announcement/>
 CC or send an email to license@isb-sib.ch).
 CC -----
 CC EMBL; X02309; CAA26179.1; --
 CC EMBL; A01157; CAA00136.1; --

DR InterPro; IPR000073; A/b hydrolase.
 DR InterPro; IPR000734; Lipase.
 DR InterPro; IPR000379; Ser. esters. site.
 DR Pfam; PF004083; abhydrolipase; 1.
 DR Pfam; PF00561; abhydrolipase; 1.
 DR PROSITE; PS00120; LIPASE SER; 1.
 KW Hydrolase; Lipid degradation; Glycoprotein; Signal.
 FT SIGNAL 1 18
 FT CHAIN 19 395 TRIACYLGLYCEROL LIPASE, LINGUAL.
 FT ACT_SITE 171 171 CHARGE RELAY SYSTEM (BY SIMILARITY).
 FT ACT_SITE 342 342 CHARGE RELAY SYSTEM (BY SIMILARITY).
 FT ACT_SITE 371 371 CHARGE RELAY SYSTEM (BY SIMILARITY).
 FT DISULFID 245 254 BY SIMILARITY.
 FT CARBOHYD 33 33 N-LINKED (GLCNAC. . .) (POTENTIAL).
 FT CARBOHYD 68 68 N-LINKED (GLCNAC. . .) (POTENTIAL).
 FT CARBOHYD 98 98 N-LINKED (GLCNAC. . .) (POTENTIAL).
 FT CARBOHYD 270 270 N-LINKED (GLCNAC. . .) (POTENTIAL).
 SQ SEQUENCE 395 AA; 44588 MW; E601854A923522EA CRC64;
 Query Match 55.9%; Score 1161; DB 1; Length 395;
 Best Local Similarity 57.7%; Pred. No. 1.8e-95;
 Matches 213; Conservative 63; Mismatches 93; Indels 0; Gaps 0;
 QY 9 DPAPFMNISEIIIOHGYPCCEYEYVATEDGYILSVNRIPLRGVLPKKTGSRPVVLLQGLV 68
 DB 27 NPEANMNIISQITWGYPCCEYEYVATEDGYILSVNRIPLRGVLPKKTGSRPVVLLQGLV 86
 QY 69 GGSNWTISNIPNNLSGLPILADAGFVMMGNSRGNWSRKHKLISIDQDFWAFSYDEMAR 128
 DB 87 ASATNWIANLPNNLSGLPILADAGFVMMGNSRGNWSRKHKLISIDQDFWAFSYDEMAR 146
 QY 129 FDLPAVINFIQKQEKIYVGVYSGQTTMGTFAPSMPELAQIKMYFALAPVATVKA 188
 DB 147 YDLPATINFIQKQEKIYVGVYSGQTTMGTFAPSMPELAQIKMYFALAPVATVKT 206
 QY 189 KSPGTKEILLPDMIMKGLFGKGFYQTRFRLQVLYLQGVILQICSNIMLLGGFNT 248
 DB 207 QSLPKLSIFPTFLFKLMFGKMFPHYTPDDFLGTEVCSREVDLLCSNLTIFPGDFDK 266
 QY 249 NNMMSRASVYAAHTLACTSVQNLHWSQAVNSGELRAPDWSGSETKLEKCNQPTPVRYR 308
 DB 267 KNLNVSRFDVYLGHNPAGTSVQDFLHWAQLVRSKTFQAFNMGSPSQNMLHYNOKTPEYD 326
 QY 309 VRDMVTPTAMWTGQDNLSPEDVLMSEVTLNLYHKNIPWAHVDFIWLDPAPHRMYN 368
 DB 327 VSMTVPVAVWNGNDILADPQDVAMLLPKLSNLFHKEILAYNHLDFIWMADAPQEVYN 386
 QY 369 EIIHLMQOE 377
 DB 387 EMISMAED 395
 RESULT 5
 LIPG CANFA
 ID LIPG CANFA STANDARD; PRT; 398 AA.
 AC P80035; O02857;
 DT 01-NOV-1991 (Rel. 20, Created)
 DT 15-JUL-1998 (Rel. 36, Last sequence update)
 DT 15-SEP-2003 (Rel. 42, Last annotation update)
 DE Triacylglycerol lipase, gastric precursor (EC 3.1.1.3) (Gastric lipase) (GL).
 GN LiPP.
 OS Canis familiaris (Dog).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Mammalia; Eutheria; Carnivora; Fissipedia; Canidae; Canis.
 OX NCBI_TaxID=9615;
 RN [1]
 RP SEQUENCE FROM N.A.
 RX TISSUE=Stomach;
 RX MEDLINE=99450174; Pubmed=10520456;
 RA Vaganay S., Joliff G., Bertaux O., Toselli E., Devignes M.D.,
 RA Benicourt C.;

DR	GO:	GO:0004806;	F:triacylglycerol lipase activity; TAS.
DR	GO:	GO:0006641;	P:triacylglycerol metabolism; NAS.
DR	InterPro:	IPR000073;	A/b hydrolase.
DR	InterPro:	IPR000734;	Lipase.
DR	InterPro:	IPR000379;	Ser esters site.
DR	Pfam:	PF04083;	abhydro_lipase; 1.
DR	Pfam:	PF00561;	abhydrolase; 1.
DR	PROSITE:	PS00120;	LIPASE_SER; 1.
KW	Hydrolase;	Lipid degradation;	Glycoprotein; Signal; Polymorphism; 3D-structure.
FT	CHAIN	1	19
FT	ACT_SITE	20	398
FT	ACT_SITE	172	172
FT	ACT_SITE	343	343
FT	ACT_SITE	372	372
FT	DISULFID	246	255
FT	CARBOHYD	34	34
FT	CARBOHYD	99	99
FT	CARBOHYD	271	271
FT	CARBOHYD	327	327
FT	VARIANT	161	161
FT	SEQUENCE	398 AA;	45237 MW; CD3BEI62LC014F0F CRC64;
QY	Query Match	53.4%;	Score 1109; DB 1; Length 398;
Db	Best Local Similarity	54.7%;	Pred. No. 7.4e-91;
QY	Matches 202;	Conservative 69;	Mismatches 98; Indels 0; Gaps 0;
QY	10	PEAFNWISEIIIOHGYPCSEYEYVATGEGYILSVNRIPRGIVOPKXTGSRPVVLLOHGLVG	69
Db	29	FEVTWNISQMITYNGYFNSEYEVVTDGVILEVNRIPYKKXSGNTGQRVPVFLOHGLLA	88
QY	70	GASNWISNLPNNSLGFILADAGFDVMGNRSRGVANSRKHTLSIQDDFFWFAPSYDEMARF	129
Db	89	SATNWSINLPNNSLAFLADAGDYDWLGNSRGNTWARRNLYYSPDSVEFWAFSPDMAKY	148
QY	130	DLPVAVINFILOKTGEKIYYGVSGQTGWGFIAPSTMPELAOKIKMYPALAPIATVKAK	189
Db	149	DLEPATPIDIVKTKQQLHVGHSGQTTTGFIAPSTNPSLAKRIKTFVALAPVATVKYTK	208
QY	190	SPGTKEFLALPDMMITKLGLFKKEFLYQTRFURLQIVYLCOGITLDICSNIMLLLGCFNTN	249
Db	209	SLINKLRVRPOSIFKPIFGDKIYPFHNFDFQFLATEVCSEMLNLLCSNALFIICGFDSK	268
QY	250	NNMGRSASYAAHTLAGTSYONILHMSQAVNSGELRAFDWGSETKNLSKCNOPTPVRVTV	309
Db	269	NENTSRLDVYLSHNPAFSGTSYQNMFMTQAYKSGKFAQYDWGSFPVNRMHEYDQSPPYXNV	328
QY	310	RDMTPTAMWTGGODMLSNEDVKMLISEVTNLIYHKNIPEWAHVDFTWGLDAPHRMVNE	369
Db	329	TANNVPVIAVWNGGXDLADPDQVCLLLPKPLNLIYHKIEIPTPHLLDFIWMADAPQEVIND	388
QY	370	IIHLMOQE 378	:
Db	389	IYSWISEDK 397	:
RESULT 8	LIP3 DROME	STANDARD;	PRT; 394 AA.
ID	LIP3 DROME	AC	O46108;
DT	16-OCT-2001	(Rel. 40,	Created)
DT	16-OCT-2001	(Rel. 40,	Last sequence update)
DT	16-OCT-2001	(Rel. 40,	Last annotation update)
DE	Lipase 3 precursor	(EC 3.1.1.-)	(Dmlip3).
GN	Lip3 OR CG8823.		
OS	Drosophila melanogaster	(Fruit fly).	
OC	Eukaryota;	Melazoa;	Arthropoda; Insecta; Pterygota;
OC	Neoptera;	Endopterygota;	Diptera; Brachycera; Muscomorpha;
OC	Ephydridae;	Drosophilidae;	Drosophila.
OX	NCBI_TaxId=7227;		
RP	SEQUENCE FROM N.A.		

STRAIN-Canton-S;
MEDLINE=98227315; PubMed=9566193;
Pistillo D., Manzi A., Tino A., Pilo Boyl P., Graziani F., Malva C.;
"The Drosophila melanogaster lipase homologs: a gene family with
tissue and developmental specific expression.";
J. Mol. Biol. 276:877-885(1998).
[2]
SEQUENCE FROM N.A.
STRAIN-Berkeley;
MEDLINE=20196006; PubMed=10731132;
Adams M.D., Celniker S.E., Holt R.A., Hoskins R.A., Galle R.F.,
Amanatides P.G., Scherzer S.E., Li P.W., Hoskins R.A., Galle R.F.,
George R.A., Lewis S.E., Richards S., Ashburner M., Henderson S.N.,
Sutton G.G., Wortman J.R., Yandell M.D., Zhang Q., Chen L.X.,
Brandon R.C., Rogers Y.-H.C., Blazek R.G., Champagne M., Pfeiffer B.D.,
Wan K.H., Achavani A., An H.-J., Andrews-Pfannkuch C., Baldwin D.,
Ballew R.M., Basu A., Baxendale J., Bayraktaroglu L., Beasley E.M.,
Beeson K.Y., Benos P.V., Berman B.P., Bhandari D., Bolshakov S.,
Borkova D., Botchan M.R., Bouck J., Brokstein P., Brottier P.,
Burtis K.C., Busan D.A., Butler H., Cadenot E., Center A., Chandra I.,
Cherry J.M., Cawley S., Dahlke C., Davenport L.B., Davies P.,
de Pablos B., DeLecher A., Deng Z., Mays A.D., Dew I., Dietz S.M.,
Dodson K., Doup L.E., Downes M., Dugan-Rocha S., Dunkov B.C., Dunn P.,
Durbin K.J., Evangelista C.C., Ferraz C., Ferrer W.M., Fleischmann W.,
Fosler C., Gabrielian A.E., Garg N.S., Gelbart W.M., Glasser K.,
Glodek A., Gong F., Gorrell J.H., Gu Z., Guan P., Harris M.,
Harris N.L., Harvey D., Heiman T.J., Hernandez J.R., Houck J.,
Hoskins D., Houston K.A., Howland T.J., Wei M.-H., Ibegwam C.,
Jalali M., Kalush F., Karpen G.H., Ke Z., Kennison J.A., Ketchum K.A.,
Kimmel B.E., Kodira C.D., Kraft C., Kravitz S., Kulp D., Lai Z.,
Liao P., Lei Y., Levitsky A.C., Li J., Li Z., Liang Y., Lin X.,
Liu X., Mattei B., McIntosh T.C., McLeod M.P., McPherson D.,
Merkulov G., Milshina N.V., Mobarry C., Morris J., Moshrefi A.,
Mount S.M., Moy M., Murphy B., Murphy L., Muzny D.M., Nelson D.L.,
Nelson D.R., Nelson K.A., Nixon K., Nusskern D.R., Pacle J.M.,
Palazzolo M., Pittman G.S., Pan S., Pollard J., Puri V., Reese M.G.,
Reinert K., Remington K., Saunders R.D.C., Scheeler F., Shen H.,
Shue B.C., Siden-Kiamos I., Simpson M., Skupski M.P., Smith T.,
Spier E., Spradling A.C., Stapleton M., Strong R., Sun E.,
Svirskas R., Tector C., Turner R., Venter E., Wang A.H., Wang X.,
Wang Z.-Y., Wasserman D.A., Weinscock G.M., Weissenbach J.,
Williams S.M., Woodage T., Worley K.C., Wu D., Yang S., Yao Q.A.,
Ye J., Yeh R.-F., Zaveri J.S., Zhan M., Zhang G., Zhao Q., Zheng L.,
Zheng X.H., Zhong F.N., Zhong W., Zhou X., Zhu S., Zhu X., Smith H.O.,
Gibbs R.A., Myers E.W., Rubin G.M., Venter J.C.;
"The genome sequence of Drosophila melanogaster.";
Science 287:2185-2195(2000).
-!- TISSUE SPECIFICITY: FAT BODY.
-!- DEVELOPMENTAL STAGE: ONLY AT LARVAL STAGES.
-!- SIMILARITY: PARTIAL WITH OTHER LIPASES (PANCREATIC, GASTRIC,
HEPATIC, LINGUAL, LIPOPROTEIN, BACTERIAL, ETC.). ALSO SIMILAR TO
LEPIDOPTERAN EGG-SPECIFIC AND YOLK PROTEINS.

This SWISS-PROT entry is copyright. It is produced through a collaboration
between the Swiss Institute of Bioinformatics and the EMBL outstation -
the European Bioinformatics Institute. There are no restrictions on its
use by non-profit institutions as long as its content is in no way
modified and this statement is not removed. Usage by and for commercial
entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
or send an email to license@isb-sib.ch).

EMBL; Y14367; CAA74737.1; -;
EMBL; AE003699; XAF54935.1; -;
FlyBase; FBgn0023495; Lip3.
InterPro; IPR000073; A/b hydrolase.
InterPro; IPR000734; Lipase.
InterPro; IPR000379; Ser esterase site.
Pfam; PF04083; abhydrolipase; 1.
Pfam; PF00561; abhydrolase; 1.
PROSITE; PS00120; LIPASE SER; 1.
HydroLase; Lipid degradation; Signal; Glycoprotein.
SIGNAL 1 20 POTENTIAL.

RA Balles R.M., Basu A., Baxendale J., Bayraktaroglu L., Beasley E.M.,
RA Beson K.Y., Benos P.V., Berman B.P., Bhandari D., Bolshakov S.,
RA Borkova D., Botchan M.R., Bouck J., Brokstein P., Brotter P.,
RA Burtis K.C., Busam D.A., Butler H., Cadieu E., Center A., Chandra I.,
RA Cherry J.M., Cawley S., Dahlke C., Davenport L.B., Davies P.,
RA de Pablos B., Delcher A., Deng Z., Mays A.D., Dew I., Dietz S.M.,
RA Dodson K.J., Doup L.E., Downes M., Duggan-Rocha S., Dunkov B.C., Dunn P.,
RA Durbin K.J., Evangelista C.C., Ferraz C., Fertiera S., Fleischmann W.,
RA Foster C., Gabriellian A.E., Garg N.S., Gelbart W.M., Glasser K.,
RA Glodok A., Gong F., Gorrell J.H., Gu Z., Guan P., Harris M.,
RA Harris N.L., Harvey D., Helman T.J., Hernandez J.R., Houck J.,
RA Hostin D., Houston K.A., Howland T.J., Wei M.-H., Ibegwam C.,
RA Jallali M., Kalush F., Karpen G.H., Ke Z., Kernison J.A., Ketchum K.A.,
RA Kimmel B.E., Kodira C.D., Kraft C., Kravitz S., Kulp D., Lai Z.,
RA Laoko P., Lei Y., Levitsky A.A., Li J., Li Z., Liang Y., Lin X.,
RA Liu X., Mattei B., McIntosh T.C., McLeod M.P., McPherson D.,
RA Merkulov G., Milshina N.V., Mobarry C., Morris J., Moshrefi A.,
RA Mount S.M., Moy M., Murphy B., Murphy L., Muzny D.M., Nelson D.L.,
RA Nelson D.R., Nelson K.A., Nixon K., Nuskern D.R., Pacleb J.M.,
RA Palazzolo M., Pittman G.S., Pan S., Pollard J., Puri V., Reese M.G.,
RA Reinert K., Remington K., Saunders R.D.C., Scheeler F., Shen H.,
RA Shue B.C., Siden-Kiamos I., Simpson M., Skupski M.P., Smith T.,
RA Spier E., Spradling A.C., Stapleton M., Strong R., Sun E.,
RA Svarksas R., Tector C., Turner R., Venter E., Wang A.H., Wang X.,
RA Wang Z.-Y., Wasserman D.A., Weinstock G.M., Weissenbach J.,
RA Williams S.M., Woodage T., Worley K.C., Wu D., Yang S., Yao Q.A.,
RA Ye J., Yeh R.-F., Zaveri J.S., Zhan M., Zhang G., Zhao Q., Zheng L.,
RA Zheng X.H., Zhong F.N., Zhong W., Zhou X., Zhu S., Zhu X., Smith H.O.,
RA Gibbs R.A., Myers E.W., Rubin G.M., Venter J.C.,
RA "The genome sequence of *Drosophila melanogaster*,"
RL Science 287:2185-2195 (2000).
RN [3]
RP SEQUENCE FROM N.A.
RC STRAIN=Berkley; TISSUE=Embryo;
RX MEDLINE=22426066; PubMed=12537569;
RA Stapleton M., Carlson J.W., Brokstein P., Yu C., Champe M.,
RA George R.A., Guarin H., Kronmiller B., Pacleb J.M., Park S., Wan K.H.,
RA Rubin G.M., Gelniker S.E.;
RT "A *Drosophila* full-length cDNA resource,";
RL Genome Biol. 3:RESEARCH0080.1-RESEARCH0080.8 (2002).
CC -!- FUNCTION: COULD BE A DIGESTIVE ENZYME.
CC -!- SUBCELLULAR LOCATION: Secreted (Probable).
CC -!- TISSUE SPECIFICITY: In 14 hour embryos expression is seen in the
CC foregut/midgut boundary.
CC -!- DEVELOPMENTAL STAGE: Expressed from 14 hour embryos through to
CC adulthood. There is a weak maternal contribution to early embryos.
CC -!- SIMILARITY: PARTIAL WITH OTHER LIPASES (PANCREATIC, GASTRIC,
CC HEPATIC, LINGUAL, LIPOPROTEIN, BACTERIAL, ETC.). ALSO SIMILAR TO
CC LEPIDOPTERAN EGG-SPECIFIC AND YOLK PROTEINS.
CC -!- CAUTION: It is uncertain whether Met-1 or Met-7 is the initiator.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
CC or send an email to license@isb-sib.ch).
CC -----
DR EMBL; Y14366; CAAY74736.1; ALT_INIT.
DR EMBL; A5003629; AAF52994.1; -.
DR EMBL; AY075506; AAL68315.1; -.
DR FlyBase; FBgn0023496; Lip1.
DR InterPro; IPR000073; A/b hydrolase.
DR InterPro; IPR000734; Lipase.
DR InterPro; IPR000379; Ser esters site.
DR Pfam; PF04083; abhydrolipase_1.
DR Pfam; PF00561; abhydrolase; 1.
DR PROSITE; PS00120; LIPASE SER; FALSE NEG.
KW Hydrolase; Lipid degradation; Signal; Glycoprotein.
FT SIGNAL 1 24 POTENTIAL.
FT CHAIN 25 439 LIPASE 1.
FT DOMAIN 30 44 POLY-GLU.

FT ACT SITE 197 CHARGE RELAY SYSTEM (BY SIMILARITY).
FT ACT SITE 393 CHARGE RELAY SYSTEM (BY SIMILARITY).
FT CARBOHYD 124 N-LINKED (GLCNAC.) (POTENTIAL).
FT CARBOHYD 151 N-LINKED (GLCNAC.) (POTENTIAL).
FT CARBOHYD 346 N-LINKED (GLCNAC.) (POTENTIAL).
FT CARBOHYD 379 N-LINKED (GLCNAC.) (POTENTIAL).
FT CARBOHYD 426 N-LINKED (GLCNAC.) (POTENTIAL).
FT CONFLICT 10 L -> I (IN REF. 1).
FT CONFLICT 213 Y -> F (IN REF. 1).
FT CONFLICT 412 Q -> E (IN REF. 1).
SQ SEQUENCE 439 AA; 50660 MW; 9E32B20BEAE93E3F CRC64;
Query Match 26.3%; Score 546; DB 1; Length 439;
Best Local Similarity 31.8%; Pred. No. 8.6e-41;
Matches 123; Conservative 83; Mismatches 157; Indels 24; Gaps 7;
QY 6 KAVDPEAFNISEIIHQGYCEVEVATEDGYILSVNRIPLGLVQPKKTSRPPVILLOH 65
DB 58 KNIKODSTLSVDKLIATKYGESEVHVTTEDGYILTWRI-----RKQA-PPFLLLOH 109
QY 66 GLVGASWISNLPNNISGLFDADAGPDVVMGNSRGNMWSRKHKTLISIDQDEFWAFSYDE 125
DB 110 GLVDSSAGFVVMGPNVSLAYLLADHNYDVMLGNARGNRYSRNHTTLDPDDESKFDFSMHE 169
QY 126 MARFPLPAVINFILOKTOEKLIVYGYSGTGTGFIAPSTPELAQKIWKYFALAPATV 185
DB 170 IGMVDLPAMIDHVLKVTGPFKLVHAGHSQGTCTFFVCMSEMPAYNDKRVSMQALAPAVYA 229
QY 186 KHAKS-PGTKFLLLPDMMIKGLFGKKEFLYQTRFLQRLVILYLCQVILDDQCSIMILLG 244
DB 230 KETEDHPYIRALSILYFNSLVGSSIREMFGFRFLCRMT-----EETERLCIEAVFGIV 283
QY 245 GNTNNMMSRASVYAHTLAGTSVQNLHWSQVNSGELAPFWSGSETKLEKCNQPTP 304
DB 284 GNNWNEFRKMPFVILGHYPAGVAAKQVYKHFQIKSGRFAPYSY-SNNKMQLYRDHLP 342
QY 305 VYRYVRDMVTPTAMWTGGQDWLSNEDVYKLLSEVTLNLYHKNP--EWAHVDFIWLGLDA 362
DB 343 PRYNLSLVTVPFVYVYSTNDLLCHPQVDSMCDLGNVTGKLYPQKEFNEMDFLWIDV 402
QY 363 PHRMVNEIHLM-----QQEETNLQ 383
DB 403 RKMLYRMLQVLGKVPFGSPSEANRSR 429
RESULT 10
TG1_YEAST STANDARD; PRT; 548 AA.
ID TG1_YEAST
AC P34163;
DT 01-FEB-1994 (Rel. 28, Created)
DT 01-FEB-1994 (Rel. 28, Last sequence update)
DT 28-FEB-2003 (Rel. 41, Last annotation update)
DE Triglyceride lipase-cholesterol esterase (EC 3.1.1.-).
GN TG1L OR YK140W OR YK15.
OS Saccharomyces cerevisiae (Baker's yeast).
CC Eukaryota; Fungi; Ascomycota; Saccharomycotina; Saccharomycetes;
CC Saccharomycetales; Saccharomycetaceae; Saccharomycetes.
CX NCBI_TaxID=4932;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=92245761; PubMed=1574929;
RA Abraham P.R., Mulder A., Van't Riet J., Planta R.J., Raue H.A.;
RT "Molecular cloning and physical analysis of an 8.2 kb segment of
RT chromosome XI of *Saccharomyces cerevisiae* reveals five tightly linked
RT genes,";
RL Yeast 8:227-238 (1992).
RN [2]
RP SEQUENCE FROM N.A.
RA Rad M.R., Xu G., Kirchrath L., Fritz C., Keuchel H., Hollenberg C.P.;
RT Submitted (MAR-1994) to the EMBL/GenBank/DDJ databases.
CC -!- SIMILARITY: Belongs to the AB hydrolase superfamily.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration

CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@isb-sib.ch).

CC EMBL; 225464; CRA80958.1; -
 CC EMBL; 228140; CRA81981.1; -
 DR PIR; S37969; S37969.
 DR SGD; S0001623; TGL1.
 DR GO; GO:000581; C:lipid particle; IDA.
 DR InterPro; IPR000073; A/B_hydrolase.
 DR InterPro; IPR000379; Ser_eatr site.
 DR Pfam; PF04083; abhydro_lipase; 1.
 DR Pfam; PF00561; abhydro_lipase; 1.
 DR KW Hydrolase; Lipid degradation.
 SQ SEQUENCE 548 AA; 62979 MW; 32D1F230701CB083 CRC64;

Query Match 20.4%; Score 422.5; DB 1; Length 548;
 Best Local Similarity 31.6%; Pred. No. 1.1e-29;
 Matches 119; Conservative 66; Mismatches 127; Indels 65; Gaps 15;
 QY 28 BEYEVEDGVLNRPVGRVLP--KKTGSRPVLLQHLGVLGASWISNLP--NNSLG 84
 DB EDHLVETEDNVLILRIP-----PISKRRFNKVVYLHGLLWCSVDWCNHERKNLP 132
 QY 85 FILDAGFDVWNGSRNAGSRKHTLSIDODEFWAFSDYDMPDLPVAVINFILOKTGQ 144
 DB FVLHDLGVWVWNGRNGKYSTAHLNPKPKSKWDFSIDFAFPDIPNSIERFIDITKV 192
 QY 145 EKIYYVGYSGTWTGCTAFSTWPELAKIKMYFALAPATVK--HAK-----SGTK 194
 DB DKVICIGFSQSGAQMFAFSLSEKLNKRVSHFIAIAPAMTPKGLHNRIVDTLAKSPGFM 252
 QY 195 FLLPLDMIMKGLFGKPELYOT---RFLQLVYLQGVILDOICSNIMLLGFGFTNN 250
 DB YL-----FFGRKIVLPSAVIWRQLTHPTLFLNLC-----ID-IANKIL-----FNWKS 293
 QY 251 MNW-----SRASVYAAHTLAGTSVQNLHWSQAVNSGELRAPD-----WGSETKNLEK 298
 DB FNILPRQKIASYAKLYST-----TSVKSIVHWFQILRSQKQFQMFESDNMLSLTREYQI 348
 QY 299 CNOQTPVRYRVRDVTPTAMTGTGQDMLSNPDEVDKMLLSEVTLNLIYKNTPEVAHVDVFIW 358
 DB 349 ANFPRT-----NIKIPILLYGGIDSLVDVWKKNLPP--FNSVFDVKVDNTEHDLIW 401
 QY 359 GLDAPHMYNEIILMQ 375
 DB 402 GKADATLVIAKVLRFIE 418

RESULT 11
 YD94_METJA STANDARD; PRT; 987 AA.
 ID YD94_METJA
 AC Q58789;
 DT 15-JUL-1998 (Rel. 36, Created)
 DT 15-JUL-1998 (Rel. 36, Last sequence update)
 DT 28-FEB-2003 (Rel. 41, Last annotation update)
 DE Hypothetical protein MJ1394.
 GN MJ1394.
 OS Methanococcus jannaschii.
 OC Archaea; Euryarchaeota; Methanococci; Methanococcales;
 OC Methanocaldococcaceae; Methanocaldococcus.
 OX NCBI_TaxID=2190;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN=JAL-1 / DSM 2661 / ATCC 43067;
 RX MEDLINE=96337999; PubMed=688087;
 RA Bult C.J., White O., Olsen G.J., Zhou L., Fleischmann R.D.,
 RA Sutton G.G., Blake J.A., FitzGerald L.M., Clayton R.A., Gocayne J.D.,
 RA Kerlavage A.R., Dougherty B.A., Tomb J.-F., Adams M.D., Reich C.I.,
 RA Overbeek R., Kirkness E.F., Weinstock K.G., Merrick J.M., Glodek A.,

RA Scott J.L., Geoghegan N.S.M., Weigman J.P., Fuhrmann J.L., Nguyen D.,
 RA Utterback T.R., Kelley J.M., Peterson J.D., Sadow P.W., Hanna M.C.,
 RA Cotton M.D., Roberts K.M., Hurst M.A., Kaine B.P., Borodovsky M.,
 RA Klenk H.-P., Fraser C.M., Smith H.O., Woese C.R., Venter J.C.;
 RT "Complete genome sequence of the methanogenic archaeon, Methanococcus
 jannaschii.";
 RL Science 273:1058-1073 (1996).
 CC -1- SUBCELLULAR LOCATION: Integral membrane protein (Potential).
 CC -1- SIMILARITY: TO M.JANNASCHII MJ1393 AND A.FULGIDUS AP2028.

CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@isb-sib.ch).

CC EMBL; U67579; AAB99404.1; -
 DR PIR; A64474; A64474.
 DR TIGR; MJ1394; -
 KW Hypothetical protein; Transmembrane; Complete proteome.
 FT TRANSMEM 12 32 POTENTIAL.
 FT TRANSMEM 958 978 POTENTIAL.
 SQ SEQUENCE 987 AA; 112360 MW; DIE628FEB28CA86D CRC64;

Query Match 5.4%; Score 111.5; DB 1; Length 987;
 Best Local Similarity 24.1%; Pred. No. 0.1;
 Matches 93; Conservative 46; Mismatches 124; Indels 123; Gaps 27;

QY 17 SRIHQHGYPCCEEVEVATEDGYILSVNRIPRG--LVQPKTGRPVVLLQHLGVLGASN 73
 DB 408 SSIQSHGNSYLEKYGIST-----SLANDPFGYKLPKEIG-RDVI--SGWYRPSN 458
 QY 74 WISNLPNNSLGFILADAGFDVWNGSRNAGSRKHT--LSIDQ-----DEFWA 120
 DB 459 W-GGGPIDRIG--LEDENFD-----GVSEFVHYSNYSIDRATNGNPTSEISPEVYN 508
 QY 121 PSYDEMARFDLPVAVINFILOKTGQEKIYVGYSGTTFGFIATFSTMPLEAKIKMYFALA 180
 DB 509 PPEDEWYFEL-----KI-----YSNGT-----ITFST-----YYONG 536
 QY 181 PT-ATVXGIAKSPGTFK-----LLLPDMNMGK-----LFGKKEFLYQTRFLRLQVLY 225
 DB 537 SLAATVSTIDNTYTFDRVHVGIVGVYVDDLEVNSKNFDFYGDKNWKL-----586
 QY 226 LCGVILDOICSNIMLLGG-----FNTNNMNSRASVYAAHTL-----AGTSVQNIL- 273
 DB 587 ---EITSANSSSEGTAVLFDGDFYKDYNTSNLN---AINWNTNITLWNSDNLVFNVLG 640
 QY 274 HWSQAVNSGELRAFWDGSETKNLEKCNQPTFVR--YRVRDVTPTAMWTGQ--DWLSN- 328
 DB 641 NYSYSERDNLAKYGFAPKILFNYNGTNTNTSIKGYVASGSYSISTDGHGTTCGEINWIVN 700
 QY 329 --PEDVKMLLSEVTLNLIYHONPEWA 352
 DB 701 TFKNDKSYSEFNLTLN-----NI--WA 720

RESULT 12
 ORC2_CABEL STANDARD; PRT; 430 AA.
 ID ORC2_CABEL
 AC Q21037;
 DT 01-NOV-1997 (Rel. 35, Created)
 DT 15-JUL-1998 (Rel. 36, Last sequence update)
 DT 28-FEB-2003 (Rel. 41, Last annotation update)
 DE Origin recognition complex subunit 2 (Ceorc2).
 GN ORC-2 OR F59E10.1.
 OS Caenorhabditis elegans.
 OC Eukaryota; Metazoa; Nematoda; Chromadorea; Rhabditida; Rhabditoidea;
 OC Rhabditidae; Peloderinae; Caenorhabditis.
 OX NCBI_TaxID=6239;
 RN [1]

RP SEQUENCE FROM N.A.
RX MEDLINE=96099401; PubMed=7502077;
RA Gavin K.A., Hidaka M., Stillman B.;
RT "Conserved initiator proteins in eukaryotes."; Science 270:1667-1671(1995).
RL [2]
RN SEQUENCE FROM N.A.
RP STRAIN=Bristol N2;
RA Swinburne J.;
RL Submitted (AUG-1994) to the EMBL/GenBank/DBJ databases.
RN REVISIONS.
RP Jones S.J.M.;
RA Submitted (MAR-1997) to the EMBL/GenBank/DBJ databases.
RL -1- FUNCTION: COMPONENT OF THE ORIGIN RECOGNITION COMPLEX (ORC) THAT BINDS ORIGINS OF REPLICATION. IT HAS A ROLE IN BOTH CHROMOSOMAL REPLICATION AND MATING TYPE TRANSCRIPTIONAL SILENCING. BINDS TO THE ARS CONSENSUS SEQUENCE (ACS) OF ORIGINS OF REPLICATION IN AN ATP-DEPENDENT MANNER.
CC -1- SUBUNIT: ORC IS COMPOSED OF SIX SUBUNITS (BY SIMILARITY).
CC -1- SUBCELLULAR LOCATION: Nuclear.
CC -1- SIMILARITY: BELONGS TO THE ORC2 FAMILY.
CC
CC This SWISS-PROT entry is copyright. It is produced through a collaboration between the Swiss Institute of Bioinformatics and the EMBL outstation - the European Bioinformatics Institute. There are no restrictions on its use by non-profit institutions as long as its content is in no way modified and this statement is not removed. Usage by and for commercial entities requires a license agreement (See <http://www.isb-sib.ch/announce/> or send an email to license@isb-sib.ch).
CC
CC EMBL; U40270; AAC46954.1; --
DR EMBL; Z36949; CA85415.1; --
DR PIR; T23001; T23001.
DR WormPep; F59E10.1; CE11492.
DR Pfam; PF04084; ORC2; 1.
KW DNA replication; Nuclear protein.
SQ SEQUENCE 430 AA; 49319 MW; 54FC086B4AD9670 CRC64;

Query Match 5.1%; Score 105; DB 1; Length 430;
Best Local Similarity 21.9%; Pred. No. 0.12;
Matches 61; Conservative 42; Mismatches 87; Indels 88; Gaps 13;

QY 52 PKTGTSPVVLQHLVGGASNWSLPPNN-----SLGF-----TLADAGFDVM--MGN 98
DB 25 PEKEGR-----QKKTNGKNAASRLQSNLEEDLEQLGFEDETVSMQSAIENYFMQK 78
QY 99 S-----RGNAWSRKH-----TSLSDQDFWAFSYDEMARFIDLPAVINILKGTGOKTY 148
DB 79 SASRMNNAWSRRRAGNNTBEIEDDEISNAITDPTKCDLPGUNVITKDNTEPEK 138
QY 149 YVGYSQGTGTFIAFTMPPELAQIKMYFAPALPIATVYHAKSPGTFKLLPDMWINGLFG 208
DB 139 RLEHLADNDFG-----KWKLYLAAG-----FNILLHGVGS 168
QY 209 KSELYOTRPLQVILQGVILQICSNIMLLGGFTN---NNMSEASVYAHTLA 265
DB 169 KRDLV---TERENELSDTYRVRDARKDGLNVKLLGAINENMKLNCVRG-----217
QY 266 GTSVQNTLHWSQV---NSGEL-----RAPDWGSE 292
DB 218 ---QSTISWARSIRRMNSQQLIILIIDNEAPDWRS 251

RESULT 13
PAPA CAVPO
ID PAPA CAVPO STANDARD; PRT; 436 AA.
AC P70683;
DT 01-NOV-1997 (Rel. 35, Created)
DT 01-NOV-1997 (Rel. 35, Last sequence update)
DT 28-FEB-2003 (Rel. 41, Last annotation update)
DE Platelet-activating factor acetylhydrolase precursor (BC 3.1.1.47)
DE (PAF acetylhydrolase) (PAF 2-acylhydrolase) (LDL-associated

phospholipase A2) (LDL-PLA(2)) (2-acetyl-1-alkylglycerophosphocholine esterase) (1-alkyl-2-acetyl-glycerophosphocholine esterase).
PLA2G7 OR PAFAH.
OS Cavia porcellus (Guinea pig).
CC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Rodentia; Hystricognathi; Caviidae; Cavia.
RN NCBI_TaxID=10141;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=Hartley; TISSUE=Liver;
RX MEDLINE=97103479; PubMed=8947850;
RA Karasawa K., Kuge O., Kawasaki K., Nishijima M., Nakano Y., Tomita M., Yokoyama K., Setaka M., Nojima S.;
RT "Cloning, expression and characterization of plasma platelet-activating factor-acetylhydrolase from guinea pig."; J. Biochem. 120:1838-844(1996).
RL -1- FUNCTION: MODULATES THE ACTION OF PLATELET-ACTIVATING FACTOR (PAF) BY HYDROLYZING THE SN-2 ESTER BOND TO YIELD THE BIOLOGICALLY INACTIVE LYSO-PAF. HAS A SPECIFICITY FOR SUBSTRATES WITH A SHORT RESIDUE AT THE SN-2 POSITION. IT IS INACTIVE AGAINST LONG-CHAIN PHOSPHOLIPIDS.
CC -1- CATALYTIC ACTIVITY: 2-acetyl-1-alkyl-sn-glycero-3-phosphocholine + H(2)O = 1-alkyl-sn-glycero-3-phosphocholine + acetate.
CC -1- SUBCELLULAR LOCATION: Extracellular.
CC -1- TISSUE SPECIFICITY: Plasma.
CC -1- SIMILARITY: BELONGS TO THE ABHYDROLASE SUPERFAMILY. LIPASE FAMILY.
CC
CC This SWISS-PROT entry is copyright. It is produced through a collaboration between the Swiss Institute of Bioinformatics and the EMBL outstation - the European Bioinformatics Institute. There are no restrictions on its use by non-profit institutions as long as its content is in no way modified and this statement is not removed. Usage by and for commercial entities requires a license agreement (See <http://www.isb-sib.ch/announce/> or send an email to license@isb-sib.ch).
CC
CC EMBL; D67037; BAA11054.1; --
DR PIR; JC5021; JC5021.
DR InterPro; IPR000734; Lipase.
DR InterPro; IPR005065; PAF-AH_P.II.
DR InterPro; IPR000379; Ser esterase.
DR Pfam; PF03403; PAF-AH_P.II.1.
DR PROSITE; PS00120; LIPASE_SER; 1.
KW Hydrolase; Lipid degradation; Glycoprotein; Signal.
FT SIGNAL 1 21
FT CHAIN 22 436
FT BY SIMILARITY.
FT PLATELET-ACTIVATING FACTOR
FT ACETYLHYDROLASE.
FT ACT_SITE 271 271
FT CHARGE_RELAY_SYSTEM (BY SIMILARITY).
FT ACT_SITE 294 294
FT CHARGE_RELAY_SYSTEM (BY SIMILARITY).
FT ACT_SITE 349 349
FT CHARGE_RELAY_SYSTEM (BY SIMILARITY).
FT CARBOHYD 76 76
FT N-LINKED (GLCNAC...) (POTENTIAL).
FT CARBOHYD 200 200
FT N-LINKED (GLCNAC...) (POTENTIAL).
FT CARBOHYD 324 324
FT N-LINKED (GLCNAC...) (POTENTIAL).
SQ SEQUENCE 436 AA; 49062 MW; C359D96E392FFE11 CRC64;

Query Match 4.7%; Score 98.5; DB 1; Length 436;
Best Local Similarity 20.5%; Pred. No. 0.48;
Matches 71; Conservative 49; Mismatches 128; Indels 99; Gaps 16;

QY 44 RIPRGLVQPKTGSR-PVLLQHLVGGASNWSLPPNNLSGLFADAGPDV-----94
DB 129 KVPKWSPLKTKGKPLIIFSHGL--GAFRSI-----YSAIGIELASHGFIIVAAVHRDE 182
QY 95 -----WMGNSRGNWSRKHKTLSIDODE-----FWAFSYDEMAR 128
DB 183 SAAATYTFQAPAAESGNRSWIYKGNLETERKQLRQRCGECQALSLLSIDEGEP 242
QY 129 FDLPAVINILQK-----TGQEKIYVYGYSQGTGTFIAFTMPPELAQIKMYFAL-----A 180
DB 243 VKNVLDLNFDDIQKGLSDRSKVAIGHSGF--GATVIQTLSE-DQFRFCGIALDPWMF 298
QY 181 PIATVYHAKSPGTFKLLPDMWINGLFGKK-EELYQTRFLRLQV-----IYL 226
DB 299 FVGSDVHSKIPQLPFFINSEYFOSANDTKKIEFYQPKERKMLVKGVSVHHNFVDTFA 358

QY 227 CQVILQICSNIMLLGGFTNNNM---SRASVAAHTLAGTSVONILHWSQAVNSGE 283
Db 359 TQKII-----GQMLSLGKIDSEVADMLINKSLAFLOKYLGLD-KNFDOWNSLME--- 408
QY 284 LRAFDWGETKLEKCNQPTFVRVVRDMVTPTAMWTCGGQDWLSNPE 330
Db 409 -----GDDENL-----IPEFTIPTTQSSGTGEORNP 436

RESULT 14
Y193 HAEIN
ID Y193 HAEIN STANDARD; PRT; 287 AA.
AC Q57427; O05013;
DT 01-NOV-1997 (Rel. 35, Created)
DT 01-NOV-1997 (Rel. 35, Last sequence update)
DT 28-FEB-2003 (Rel. 41, Last annotation update)
DE Putative esterase/lipase H10193 (EC 3.1.1.-).
CN H10193
OS Haemophilus influenzae.
OC Bacteria; Proteobacteria; Gammaproteobacteria; Pasteurellales;
OC Pasteurellaceae; Haemophilus.
OX NCBI_TaxID=727;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=RD / KW20 / ATCC 51907;
RX MEDLINE=95350630; PubMed=7542800;
RA Felschmann R.D., Adams M.D., White O., Clayton R.A., Kirkness E.F.,
RA Kerlavage A.R., Bult C.J., Tomb J.-F., Dougherty B.A., Merrick J.M.,
RA McKenney K., Sutton G., Fitzhugh W., Fields C.A., Gocayne J.D.,
RA Scott J.D., Shirley R., Liu L.-I., Glodek A., Kelley J.M.,
RA Weidman J.P., Phillips C.A., Spriggs T., Hedblom E., Cotton M.D.,
RA Utterback T.R., Hanna M.C., Nguyen D.T., Saudek D.M., Brandon R.C.,
RA Fine L.D., Fritchman J.L., Fuhrmann J.L., Geoghagen N.S.M.,
RA Gnehm C.L., McDonald L.A., Small K.V., Fraser C.M., Smith H.O.,
RA Venter J.C.;
RT "Whole-genome random sequencing and assembly of Haemophilus influenzae
Rd.";
RL Science 269:496-512 (1995).
CC -!- SIMILARITY: BELONGS TO A FAMILY OF ESTERASES THAT GROUPS TOGETHER
CC PSEUDOMONAS TROPINESTERASE, DMPD, TOOF AND XLYP.
CC
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
CC or send an email to license@isb-sib.ch).
CC
CC EMBL; U32704; AAC21862.1; -.
DR PIR; E64053; E64053.
DR TIGR; H10193; -.
DR InterPro; IPR000073; A/b hydrolase.
DR InterPro; IPR003089; AB hydrolase.
DR InterPro; IPR000379; Ser esterase.
DR Pfam; PF00561; abhydrolase; 1.
DR PRINTS; PR00111; ABHYDROLASE.
KW Hypothetical protein; Hydrolase; Serine esterase; Complete proteome.
FT ACT_SITE 119 119 BY SIMILARITY.
FT ACT_SITE 286 286 BY SIMILARITY.
SQ SEQUENCE 287 AA; 32987 MW; F2B548619C060619 CRC64;

Query Match 4.6%; Score 96.5; DB 1; Length 287;
Best Local Similarity 19.1%; Pred. No. 0.41;
Matches 62; Conservative 39; Mismatches 119; Indels 105; Gaps 12;
QY 51 QPKTGSRPVLLQGLVGAGSNWISLNNLSGLFADAGFDVWNGNSRGNWSRKHT 110
Db 39 QVKQTINTPVLIFIRHGLFGMDN-----LGV-----ARAFSEHSYI 75
QY 111 LSIDQDEFWARYSDWARDFLPAV-INPILQKTGQEKIVYGVSGTWTMGIAFTWPEL 169

Db 76 LRIDLNRHGHSHFSEKQNYQLMAEDVIAVIRHLNLSKVILIGHSMGGKTAMKITALCP 135
QY 170 AOKIKMYFALAPIATVKHAKSGTKFLLPDM-----MIKGLGKKEFLYQTR 217
Db 136 VEK-----LIVIDMSFMPYEGFGHGVFNGLFAVNAKAPEN- 171
QY 218 FLRLVYLCGVVLDQICSNIMLLGGFTNNNMNRASVAAHTLAGTSVONILHWSQ 277
Db 172 --RQAKPILKQBINDE--DVQFMLKSPDVNSDCFRNL-----TALFNNYANIMDWE- 222
QY 278 AVNSGELRAFDWGETKLEKCNQPTFVRVVRDMVTPTAMWTCGGQDWLSNPEVYKMLS 337
Db 223 -----KRVF-----TPTFLIKGNSSVIKIENSEKILE 251
QY 338 EVTNLIYH--KNIPWAHV---DFI 357
Db 252 QFNATAFITINGSCHVHAERKDFV 276

RESULT 15
PRXC STRAU
ID PRXC STRAU STANDARD; PRT; 278 AA.
AC Q31168;
DT 15-DEC-1998 (Rel. 37, Created)
DT 15-DEC-1998 (Rel. 37, Last sequence update)
DT 28-FEB-2003 (Rel. 41, Last annotation update)
DE Non-heme chloroperoxidase (EC 1.11.1.10) (Chloride peroxidase)
DE (CPO-T) (Chloroperoxidase T).
GN CPO OR CPOT.
OS Streptomyces aureofaciens.
OC Bacteria; Actinobacteria; Actinobacteridae; Actinomycetales;
OC Streptomycinae; Streptomycetaceae; Streptomyces.
OX NCBI_TaxID=1894;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=TUE24;
RA Pelletier I., Altenbuchner J., van Pee K.-H.;
RL Submitted (OCT-1997) to the EMBL/GenBank/DBJ databases.
RN [2]
RP X-RAY CRYSTALLOGRAPHY (1.75 ANGSTROMS).
RC STRAIN=TUE24;
RX MEDLINE=98307994; PubMed=9642069;
RA Hofmann B., Tolzer S., Pelletier I., Altenbuchner J., van Pee K.-H.,
RA Hecht H.-J.;
RT "Structural investigation of the cofactor-free chloroperoxidases.";
RL J. Mol. Biol. 279:889-900 (1998).
CC -!- CATALYTIC ACTIVITY: 2 RH + 2 chloride + H(2)O(2) = 2 RCL + 2
CC H(2)O.
CC -!- SUBUNIT: Homodimer.
CC -!- SIMILARITY: TO OTHER BACTERIAL NON-HEME BROMO- AND CHLORO-
CC PEROXIDASES.
CC
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
CC or send an email to license@isb-sib.ch).
CC
CC EMBL; AF031242; AAB86626.1; -.
DR PDB; 1A7U; 17-JUN-98.
DR InterPro; IPR000073; A/b hydrolase.
DR InterPro; IPR003089; AB hydrolase.
DR InterPro; IPR000639; Epox hydrolase.
DR InterPro; IPR000379; Ser esterase.
DR Pfam; PF00561; abhydrolase; 1.
DR PRINTS; PR00111; ABHYDROLASE.
KW Oxidoreductase; Peroxidase; Chloride; 3D-structure.
FT ACT_SITE 99 99 BY SIMILARITY.
FT ACT_SITE 229 229 BY SIMILARITY.

BY SIMILARITY.

FT	ACT SITE	258	258
FT	STRAND	3	9
FT	TURN	10	11
FT	STRAND	12	22
FT	STRAND	26	30
FT	TURN	33	34
FT	TURN	37	40
FT	HELIX	41	49
FT	TURN	50	51
FT	STRAND	53	57
FT	TURN	60	61
FT	TURN	63	64
FT	HELIX	74	88
FT	TURN	89	89
FT	STRAND	93	98
FT	TURN	99	99
FT	HELIX	100	112
FT	TURN	115	116
FT	STRAND	117	123
FT	STRAND	131	131
FT	TURN	137	138
FT	STRAND	141	141
FT	HELIX	143	155
FT	HELIX	157	168
FT	TURN	169	170
FT	HELIX	171	174
FT	STRAND	175	175
FT	TURN	176	178
FT	STRAND	179	179
FT	HELIX	181	193
FT	HELIX	196	205
FT	TURN	206	207
FT	TURN	211	213
FT	HELIX	214	216
FT	STRAND	221	226
FT	TURN	227	228
FT	HELIX	234	236
FT	TURN	237	237
FT	HELIX	238	244
FT	TURN	246	247
FT	STRAND	249	253
FT	TURN	254	255
FT	TURN	258	259
FT	HELIX	260	263
FT	TURN	264	264
FT	HELIX	265	278
SQ	SEQUENCE	278 AA; 30354 MW; 8C97A87251FBEDES CRC64;	
Query Match 4.5%; Score 93.5; DB 1; Length 278;			
Best Local Similarity 22.6%; Pred. No. 0.72;			
Matches 72; Conservative 36; Mismatches 114; Indels 97; Gaps 17;			
QY	58	RPVLLQHLGVLGGASNIPLNNSLGFILADAGFDVWMGNSRGNAMSRKHKTLISIDQDE	117
Db	25	QPVULI-HGFFLSGHSW-----ERQSAALLDAGYRIVYDRRGFGQSQPTT-----	70
QY	118	FWAFSYDEMARFDLPVAVINFILQKTQEKIYVGYSGQTTMGFTAFSTMPELAKIKWYF	177
Db	71	--GYDYDTFA-----ADLNTVLETLDLQDAVLGVFSMGT-----GEVARYVSSY-	112
QY	178	ALAPIATYKHAKSPGPKELLAPDMIKLPGKKEPLYQTRPLRQLVIVLCGOVILDOICS	237
Db	113	GTARIAKVAFLASL-EPFLKTKTDNPDGA-APKEF-----FDGIVA	151
QY	238	NIMLLGGFNTNMMMSRASVYAHTLAGT-----SVQNIHWSQAVNSGELRA-----	288
Db	152	AVKADRYAFYTGFFN-----DFYNLDENLGTRISEAVEN--SWNTAASGGFPFAAAAPT	205
QY	289	WGSEYK-NLEKNQPTPVRYVRDVTFTAMWTGGQDWLSNPEDVKMLLSEVTLNLIYKRN	347
Db	206	WYTFDPRADIPRIDVPALILHGTGDRTLPI-----ENTARVFHKA	244
QY	348	IP--EWAHVDFIWLGLDAPH	364

Db 245 LPSAAYVEVE-----GAPH 258

Search completed: February 19, 2004, 14:21:13
Job time : 11:1144 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2004 Compugen Ltd.

OM protein - protein search, using sw model

Run on: February 19, 2004, 14:14:34 ; Search time 34.8598 Seconds
(without alignments)
1926.040 Million cell updates/sec

Title: US-10-042-431-47

Perfect score: 2247

Sequence: 1 MLETLRSQVSHRHEMML.....IHLMQBETNLSQRCRAVL 423

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 1107863 seqs, 158726573 residues

Total number of hits satisfying chosen parameters: 1107863

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : A Geneseq 19Jun03:*

- 1: /SIDS1/gcgdata/geneseq/geneseq-emb1/AA1980.DAT.*
- 2: /SIDS1/gcgdata/geneseq/geneseq-emb1/AA1981.DAT.*
- 3: /SIDS1/gcgdata/geneseq/geneseq-emb1/AA1982.DAT.*
- 4: /SIDS1/gcgdata/geneseq/geneseq-emb1/AA1983.DAT.*
- 5: /SIDS1/gcgdata/geneseq/geneseq-emb1/AA1984.DAT.*
- 6: /SIDS1/gcgdata/geneseq/geneseq-emb1/AA1985.DAT.*
- 7: /SIDS1/gcgdata/geneseq/geneseq-emb1/AA1986.DAT.*
- 8: /SIDS1/gcgdata/geneseq/geneseq-emb1/AA1987.DAT.*
- 9: /SIDS1/gcgdata/geneseq/geneseq-emb1/AA1988.DAT.*
- 10: /SIDS1/gcgdata/geneseq/geneseq-emb1/AA1989.DAT.*
- 11: /SIDS1/gcgdata/geneseq/geneseq-emb1/AA1990.DAT.*
- 12: /SIDS1/gcgdata/geneseq/geneseq-emb1/AA1991.DAT.*
- 13: /SIDS1/gcgdata/geneseq/geneseq-emb1/AA1992.DAT.*
- 14: /SIDS1/gcgdata/geneseq/geneseq-emb1/AA1993.DAT.*
- 15: /SIDS1/gcgdata/geneseq/geneseq-emb1/AA1994.DAT.*
- 16: /SIDS1/gcgdata/geneseq/geneseq-emb1/AA1995.DAT.*
- 17: /SIDS1/gcgdata/geneseq/geneseq-emb1/AA1996.DAT.*
- 18: /SIDS1/gcgdata/geneseq/geneseq-emb1/AA1997.DAT.*
- 19: /SIDS1/gcgdata/geneseq/geneseq-emb1/AA1998.DAT.*
- 20: /SIDS1/gcgdata/geneseq/geneseq-emb1/AA1999.DAT.*
- 21: /SIDS1/gcgdata/geneseq/geneseq-emb1/AA2000.DAT.*
- 22: /SIDS1/gcgdata/geneseq/geneseq-emb1/AA2001.DAT.*
- 23: /SIDS1/gcgdata/geneseq/geneseq-emb1/AA2002.DAT.*
- 24: /SIDS1/gcgdata/geneseq/geneseq-emb1/AA2003.DAT.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	2247	100.0	423	22 AAB66065	Human TANGO 294.
2	2247	100.0	423	22 AAB66065	Amino acid sequenc
3	2174	96.8	409	22 AAE11931	Human CG162 (Or C5
4	2076	92.4	390	22 AAB66067	Human TANGO 294 ma
5	2013.5	89.6	423	23 AAB66067	Human lipid-associ
6	1991	88.6	398	23 AEG71000	Human lipase prote
7	1991	88.6	398	23 AAU99164	Human lysosomal ac
8	1536.5	68.4	371	23 AAU99165	Human lysosomal ac
9	1289	57.4	399	22 AAB90783	Human shear stress

10	1289	57.4	399	22 AAB66061	Human lysosomal ac
11	1282	57.1	392	22 AAO18227	Human lysosomal ac
12	1191.5	53.0	449	23 AAU30498	Novel human secret
13	1166	51.9	398	14 AAR37302	RGL precursor. Or
14	1165	51.8	221	22 AAB66068	Human TANGO 294 ex
15	1156.5	51.5	395	6 AAF50322	Rat lingual lipase
16	1143	50.9	395	23 AAE14744	Human triacylglyce
17	1143	50.9	399	23 AAU77496	Human lipid metabo
18	1141.5	50.8	379	15 AAR56870	Canine gastric lip
19	1141.5	50.8	379	17 AAW09382	Dog gastric lipase
20	1141.5	50.8	380	15 AAR56871	Canine gastric lip
21	1139	50.7	399	22 AAG67513	Amino acid sequenc
22	1136	50.6	397	23 AAE17308	Human lysosomal ac
23	1127	50.2	378	23 ABB76189	Kid goat pregastric
24	1119	49.8	398	7 AAP60724	Sequence of pregas
25	1119	49.8	398	7 AAP60658	Sequence of human
26	1119	49.8	398	17 AAW09383	Human gastric lipa
27	1119	49.8	398	22 AAB66086	Human lipase prote
28	1119	49.8	398	24 AAB96698	Human gastric lipa
29	1101	49.0	403	23 AAU77493	Human lipid metabo
30	1090	48.5	427	23 AAU98539	Human lysosomal ac
31	1034	46.0	401	23 AEG94641	Human NOV6b protei
32	1031	45.9	395	23 AAO18226	Human lysosomal ac
33	1031	45.9	395	24 AEG73144	Human lipase prote
34	1006	44.8	373	23 AEP53569	Human NOV1 protein
35	995.5	44.3	362	24 AAB96697	Human triacylglyce
36	977	43.5	365	23 AAE17307	Human lysosomal ac
37	909	40.5	331	24 ABR41626	Human DITHP secret
38	907	40.4	731	23 AAE14746	Human triacylglyce
39	901.5	40.1	390	23 AEG94640	Human NOV6a protei
40	783	34.8	144	22 AAB66070	Human TANGO-294 CY
41	781	34.8	276	23 AAE14745	Human triacylglyce
42	777	34.6	289	23 AAU77494	Human lipid metabo
43	745.5	33.2	280	23 AAE14743	Human triacylglyce
44	724.5	32.2	656	22 AEG26839	Novel human diagno
45	691	30.8	233	22 AAB61608	Human protein HP03

ALIGNMENTS

RESULT 1

AAB66065
ID AAB66065 standard; Protein; 423 AA.

AC AAB66065;

DT 30-MAR-2001 (first entry)

DE Human TANGO 294.

KW TANGO protein; INTERCEPT protein; neurological disorder;
KW central nervous system; focal brain disorder; bipolar affective disorder;
KW global-diffuse cerebral disorder; cerebrovascular; Alzheimer's disease;
KW senile dementia; Huntington's disease; amyotrophic lateral sclerosis;
KW Parkinson's; Gilles de la Tourette's syndrome; hypertension; sleep;
KW neuropsychiatric; psychoactive substance use; anxiety.

OS Homo sapiens.

PN WO200077239-A2.

PD 21-DEC-2000.

PF 24-MAY-2000; 2000WO-US14858.

PR 14-JUN-1999; 99US-0333159.

PA (MILL-) MILLENNIUM PHARM INC.

PI McCarthy SA, Fraser CC, Sharp JD, Barnes TM;

PP WPI; 2001-032313/04.

DR N-PSDB; AAF45131, AAF45132.
 XX TANGO and INTERCEPT nucleic acids, proteins, and antibodies, useful for
 PT screening assays and diagnostic assays and for the treatment of
 PT neurological diseases such as Alzheimer's, Parkinson's and Huntington's
 PT disease -
 XX
 PS Claim 8; Fig 6; 359pp; English.
 XX
 CC The present invention relates to TANGO or INTERCEPT proteins and coding
 CC sequences (see AAF45121-F45136 and AAF45138-F45139 and AAF6031-B66057,
 CC AAF6064-B66083 and AAF66085). The TANGO/INTERCEPT proteins and coding
 CC sequences are useful for the treatment of neurological disorders such as
 CC central nervous system (CNS) disorders, CNS-related disorders, focal
 CC brain disorders, global-diffuse cerebral disorders and other
 CC neurological and cerebrovascular disorders. The CNS disorders include
 CC Alzheimer's disease, senile dementia, Huntington's disease, amyotrophic
 CC lateral sclerosis, Parkinson's, Gilles de la Tourette's syndrome,
 CC autonomic function disorders such as hypertension and sleep disorders,
 CC neuropsychiatric disorders, psychoactive substance use disorders,
 CC anxiety, and bipolar affective disorder.
 XX
 SQ Sequence 423 AA;
 Query Match 100.0%; Score 2247; DB 22; Length 423;
 Best Local Similarity 100.0%; Pred. No. 6.8e-220;
 Matches 423; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 Qy 1 MLETLSRQWIVSHRMEMWLLILVAYMFORNVNSVHMTKAVDPEAFNISEIIHQGYPC 60
 Db 1 MLETLSRQWIVSHRMEMWLLILVAYMFORNVNSVHMTKAVDPEAFNISEIIHQGYPC 60
 Qy 61 EYEYVATEDGVILSVNRIIPRLGVOPKGTGRPVVLLQHLVGGASNWSINLPNSLGFIL 120
 Db 61 EYEYVATEDGVILSVNRIIPRLGVOPKGTGRPVVLLQHLVGGASNWSINLPNSLGFIL 120
 Qy 121 ADAGFDVWNGSRGNARSKHKLISIDQDFWAFSDYEMARFDLPVAFINFILOKTQOEKI 180
 Db 121 ADAGFDVWNGSRGNARSKHKLISIDQDFWAFSDYEMARFDLPVAFINFILOKTQOEKI 180
 Qy 181 YVGYSGQTGTFIAFTMPLEAKIKMYFALAPIATVKGAKSPGTKFLLLPDMTKGLF 240
 Db 181 YVGYSGQTGTFIAFTMPLEAKIKMYFALAPIATVKGAKSPGTKFLLLPDMTKGLF 240
 Qy 241 GKKEFLYQTRFLRQVLYLCCQVILDOICSNIMLLGENTNNMNSRASVYAAHTLAGT 300
 Db 241 GKKEFLYQTRFLRQVLYLCCQVILDOICSNIMLLGENTNNMNSRASVYAAHTLAGT 300
 Qy 301 SVQNILHWSQAVNSGELRAFDPWGSSETKNLEKCNQPTPVYRVDRMTVPTAMWTGGQDWLS 360
 Db 301 SVQNILHWSQAVNSGELRAFDPWGSSETKNLEKCNQPTPVYRVDRMTVPTAMWTGGQDWLS 360
 Qy 361 NPEDVKMLLSEVTNLIYHKNIPENAVHDFIWLGDAPHRMYNEIHLMOQEBTNSQGRCE 420
 Db 361 NPEDVKMLLSEVTNLIYHKNIPENAVHDFIWLGDAPHRMYNEIHLMOQEBTNSQGRCE 420
 Qy 421 AVL 423
 Db 421 AVL 423
 RESULT 2
 ABU08369
 ID ABU08369 standard; Protein; 423 AA.
 AC ABU08369;
 XX
 DT 03-JUN-2003 (first entry)
 DE Amino acid sequence for human TANGO 294.
 DE Human; TANGO 294; INTERCEPT; cellular process; tissue typing;
 KW forensic biology; cellular protease activity; cell interaction;
 KW

development; blood disorder; haematopoietic cell-related disorder;
 growth, cell proliferation; cell differentiation; gamma delta T-cell;
 immune system; metabolic disorder; homeostatic disorder; anaemia;
 development bone disorder; osteoporosis; bacterial infection;
 thrombocytopaenia; renal failure; liver disease; cystic fibrosis;
 obesity; cancer; bronchitis; asthma; emphysema; pulmonary oedema;
 respiratory distress syndrome; osteopathic; antibacterial; antianaemic;
 thrombolytic; nephrotropic; antiobesity; hepatotropic; cytostatic;
 antiinflammatory; antiasthmatic.
 XX Homo sapiens.
 OS
 FH Key Location/Qualifiers
 FT Peptide 1..33
 FT /label= Predicted signal peptide
 FT /note= "Given as SEQ ID No:48 and specifically
 FT claimed in Claim 8"
 FT Protein 34..423
 FT /label= Predicted mature TANGO 294 protein
 FT /note= "Given as SEQ ID No:49 and specifically
 FT claimed in Claim 8"
 FT Domain 34..254
 FT /label= Extracellular domain
 FT /note= "Given as SEQ ID No:50 and specifically
 FT claimed in Claim 8"
 FT Domain 255..279
 FT /label= Transmembrane domain
 FT /note= "Given as SEQ ID No:51 and specifically
 FT claimed in Claim 8"
 FT Domain 280..423
 FT /label= Cytoplasmic domain
 FT /note= "Given as SEQ ID No:52 and specifically
 FT claimed in Claim 8"
 XX
 PN US2002182675-A1.
 XX
 PD 05-DEC-2002.
 XX
 PF 25-OCT-2001; 2001US-0042431.
 XX
 PR 14-JUN-1999; 98US-0333159.
 PR 24-MAY-2000; 2000US-0578063.
 XX
 PA (MILL-) MILLENNIUM PHARM INC.
 XX
 PI McCarthy SA, Fraser CC, Sharp JD, Barnes TM;
 XX
 DR WPI; 2003-328617/31.
 DR N-PSDB; ABX94106.
 XX
 PT New TANGO and INTERCEPT proteins, useful as modulating agents in
 PT regulating a variety of cellular processes, in chromosome mapping, in
 PT tissue typing, and in forensic biology -
 XX
 PS Claim 8; Fig 6A-6C; 232pp; English.
 XX
 CC The present invention relates to the isolation of novel TANGO or
 CC INTERCEPT proteins, and the polynucleotide sequences encoding them.
 CC The invention discloses sequences for human TANGO 202, TANGO 234,
 CC TANGO 255, TANGO 273, TANGO 286, TANGO 294 and INTERCEPT 296. The
 CC sequences for murine TANGO 202 and TANGO 273 are also provided. The
 CC TANGO polypeptides and the polynucleotide sequences encoding them are
 CC useful as modulating agents in regulating a variety of cellular
 CC processes. The polynucleotide sequences are useful as primers or
 CC hybridisation probes for the detection of nucleic acids encoding
 CC TANGO polypeptides, chromosome mapping, tissue typing, and in
 CC forensic biology. TANGO 202 can be used to diagnose, prevent or
 CC treat disorders relating to aberrant cellular protease activity,
 CC inappropriate interaction of cells with mediators, inappropriate
 CC development, and blood and haematopoietic cell-related disorders.
 CC TANGO 234 can be used to modulate growth, proliferation, survival,
 CC differentiation, and activity of gamma delta T-cells. TANGO 265 can
 CC be used to prevent, diagnose and treat disorders characterised by

aberrant organisation or development of a tissue or organ, and for modulating differentiation of cells of the immune system. TANGO 273 is useful for diagnosing, treating or preventing e.g. metabolic, homeostatic and developmental bone disorders (e.g. osteoporosis), and bacterial infection. TANGO 286 is useful for treating or preventing e.g. anaemia, thrombocytopenia, renal failure or liver disease. TANGO 294 is useful for treating or preventing e.g. cystic fibrosis or obesity. INTERCEPT 296 is useful for diagnosing, treating or preventing e.g. cancers, bronchitis, cystic fibrosis, asthma, emphysema, pulmonary oedema, or adult and infant respiratory distress syndromes. The present sequence represents human TANGO 294.

Query Match 100.0%; Score 2247; DB 24; Length 423;
Best Local Similarity 100.0%; Pred. No. 6.8e-220;
Matches 423; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 MLETSLRWIVSHRMEMWLLILVAYMFQNNVSVHMTKAVDPFAFNISEIIHQGYPC 60
Db 1 MLETSLRWIVSHRMEMWLLILVAYMFQNNVSVHMTKAVDPFAFNISEIIHQGYPC 60

Qy 61 EYEVATDGYLSVNRIPRGLVOPKKTGSRPVVLLQHGLVGGASWISNLPNNSLGFIL 120
Db 61 EYEVATDGYLSVNRIPRGLVOPKKTGSRPVVLLQHGLVGGASWISNLPNNSLGFIL 120

Qy 121 ADAGFDVVMGNSRGNWNRKHTLSIDQDEFWAFSDYEMARFDPVAVINFILQKTQGEKI 180
Db 121 ADAGFDVVMGNSRGNWNRKHTLSIDQDEFWAFSDYEMARFDPVAVINFILQKTQGEKI 180

Qy 181 YVGVSGQTTWGTAFSTMPDLAKIKMYFALAPIATVYKAKSPGTFKLLPDMNKGFLP 240
Db 181 YVGVSGQTTWGTAFSTMPDLAKIKMYFALAPIATVYKAKSPGTFKLLPDMNKGFLP 240

Qy 241 GKGFPLYQTRFLRLQVILYLCQVILQDQCSNIMLLGCFNTNNMGRASYAHTLAGT 300
Db 241 GKGFPLYQTRFLRLQVILYLCQVILQDQCSNIMLLGCFNTNNMGRASYAHTLAGT 300

Qy 301 SVONILHWSQVNSGELRAFQWSETKNLEKCNQPTPVRYVRDMVTPTAMWTGGQDWLS 360
Db 301 SVONILHWSQVNSGELRAFQWSETKNLEKCNQPTPVRYVRDMVTPTAMWTGGQDWLS 360

Qy 361 NPEDVKMLLSEVTNLIYHKNIPWAHVDVFWGLDAPHMYNEIHLMQQEBTNLSQGRCE 420
Db 361 NPEDVKMLLSEVTNLIYHKNIPWAHVDVFWGLDAPHMYNEIHLMQQEBTNLSQGRCE 420

Qy 421 AVL 423
Db 421 AVL 423

RESULT 3

ID AAE11931 standard; Protein; 409 AA.
AC AAE11931;
XX AAE11931;
DT 18-DEC-2001 (first entry)
XX Human CG162 (or C59) lipase protein #2.
DE Human; apolipoprotein; lipase; lipid metabolism; myocardial infarction;
XX cardiovascular disease; arterial thrombosis; thrombolytic; antilipemic;
KW cerebral ischaemia; arterial thrombosis; thrombolytic; antilipemic;
KW coronary artery thrombosis; cerebral artery thrombosis; stroke;
KW intracardiac thrombosis; gene therapy; cardiovascular; vasodilator;
KW neuroprotectant; cerebroprotective.
XX Homo sapiens.
OS Homo sapiens.
XX Key Location/Qualifiers
FH Peptide 1..19
FT

Protein /label= Signal_peptide
20.409
/note= "Human mature CG162 (or C59) lipase protein"

WO200179446-A2.
25-OCT-2001.
16-APR-2001; 2001WO-US12529.
14-APR-2000; 2000US-197137P.
20-JUN-2000; 2000US-0598042.
03-AUG-2000; 2000US-0631451.
17-NOV-2000; 2000US-0667298.
17-NOV-2000; 2000US-0714936.
(HYSE-) HYSEQ INC.
PA Ballinger DG, Loeb D, Montgomery JR, Tang TY, Zhou P, Goodrich R;
PI Liu C, Asundi V, Zhao QA, Wehrman T, Drmanac RI, Ren F, Qian XB;
PI Wang D;
XX WPI; 2001-611724/70.
XX N-PSDB; AAD19226.
XX Nucleic acids encoding human apolipoproteins, lipases, and lipoprotein
XX receptor polypeptides, useful for preventing diagnosing and treating
XX lipid metabolism disorders, thrombosis and cardiovascular diseases -
XX Claim 10; Fig 3; 266pp; English.
XX The invention relates to polynucleotides encoding proteins CG122, CG179,
XX CG95, CG121, CG162, CG27, CG153 and CG168 which are related to proteins
XX involved in lipid metabolism and cardiovascular disease such as human
XX apolipoproteins, lipases and lipoprotein receptor proteins. These DNA
XX and protein sequences are useful for treating or preventing disorders
XX associated with apolipoproteins, lipases and lipoprotein receptor (ALLr)
XX expression and for treating lipid metabolism, cardiovascular diseases
XX and thrombosis. Antibodies against these proteins are useful for
XX determining the presence of or predisposition to a disease associated
XX with altered levels of these sequences. ALLr polypeptides are also
XX useful for identifying agents (agonists and antagonists) that bind to
XX them and cells expressing ALLr proteins are useful for identifying a
XX therapeutic agent for use in treatment of a pathology related to
XX aberrant expression or physiological interactions of this polypeptide.
XX Vectors comprising these DNA and protein sequences are also useful for
XX producing ALLr proteins. The nucleic acids and polypeptides of the
XX invention are also useful for the treatment of occlusive cardiovascular
XX diseases, myocardial infarction, cerebral ischaemia, angina, arterial
XX thrombosis, coronary artery thrombosis and cerebral artery thrombosis
XX or intracardiac thrombosis and stroke. The nucleotides of the invention
XX are used in gene therapy. The present sequence is human CG162 (or C59)
XX lipase protein.
XX SQ Sequence 409 AA;

Query Match 96.8%; Score 2174; DB 22; Length 409;
Best Local Similarity 100.0%; Pred. No. 1.8e-212;
Matches 409; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 15 MEMWLLILVAYMFQNNVSVHMTKAVDPFAFNISEIIHQGYPCYEEVATEGYILS 74
Db 1 MEMWLLILVAYMFQNNVSVHMTKAVDPFAFNISEIIHQGYPCYEEVATEGYILS 60

Qy 75 VNRIPRGLVOPKKTGSRPVVLLQHGLVGGASWISNLPNNSLGFILADAGFDVVMGNSRG 134
Db 61 VNRIPRGLVOPKKTGSRPVVLLQHGLVGGASWISNLPNNSLGFILADAGFDVVMGNSRG 120

Qy 135 NAWSRKHKTLSDQDEFWAFSDYEMARFDPVAVINFILQKTQGEKIYYGVSGQTTMGFI 194
Db 121 NAWSRKHKTLSDQDEFWAFSDYEMARFDPVAVINFILQKTQGEKIYYGVSGQTTMGFI 180

Qy 195 AFSTMPDLAQKIMVFPALAPIATVYKAKSPGTFKLLPDMNKGFLGKKEFLYQTRFLRQ 254

Db 181 AFSTMPPELAQKIMVYFALAPIATVYKIAKSPGTFKLLPDMTKGLFGKCFYQTRFLRQ 240
 QY 255 LVLYLGGQVILDOICSNIMLLGGFTNNMNSRASVYAAHTLAGTSVQNIHLWQAVNS 314
 Db 241 LVLYLGGQVILDOICSNIMLLGGFTNNMNSRASVYAAHTLAGTSVQNIHLWQAVNS 300
 QY 315 GELRAFDMGSETKXLEKCNQPTPVYRVRDMTPTAMWTGGQDWSNPEDVKLLSEVTN 374
 Db 301 GELRAFDMGSETKXLEKCNQPTPVYRVRDMTPTAMWTGGQDWSNPEDVKLLSEVTN 360
 QY 375 LIYHKNIPWAHVDFTWGLDAPHRMYNEIHLMOQEBTNLSQRCBAVL 423
 Db 361 LIYHKNIPWAHVDFTWGLDAPHRMYNEIHLMOQEBTNLSQRCBAVL 409

RESULT 4

AAB66067
 ID AAB66067 standard; Protein; 390 AA.
 XX
 AC AAB66067;
 XX
 DT 30-MAR-2001 (first entry)
 XX
 DE Human TANGO 294 mature protein.
 XX
 KW TANGO protein; INTERCEPT protein; neurological disorder;
 KW central nervous system; focal brain disorder; bipolar affective disorder;
 KW global-diffuse cerebral disorder; cerebrovascular; Alzheimer's disease;
 KW senile dementia; Huntington's disease; amyotrophic lateral sclerosis;
 KW Parkinson's; Gilles de la Tourette's syndrome; hypertension; sleep;
 KW neuropsychiatric; psychoactive substance use; anxiety.
 XX
 OS Homo sapiens.
 XX
 PN WO200077239-A2.
 XX
 PD 21-DEC-2000.
 XX
 PF 24-MAY-2000; 2000WO-US14958.
 XX
 PR 14-JUN-1999; 99US-0333159.
 XX
 PA (MILL-) MILLENNIUM PHARM INC.
 XX
 PI McCarthy SA, Fraser CC, Sharp JD, Barnes TW;
 XX
 DR WPI; 2001-032313/04.
 XX
 KW TANGO and INTERCEPT nucleic acids, proteins, and antibodies, useful for
 PT screening assays and diagnostic assays and for the treatment of
 PT neurological diseases such as Alzheimer's, Parkinson's and Huntington's
 PT disease -
 XX
 PS Claim 8; Pages 324-325; 359pp; English.

CC The present invention relates to TANGO or INTERCEPT proteins and coding
 CC sequences (see AAF45121-F45136 and AAF45138-F45139 and AAB66031-B66057,
 CC AAB66064-B66083 and AAB66085). The TANGO/INTERCEPT proteins and coding
 CC sequences are useful for the treatment of neurological disorders such as
 CC central nervous system (CNS) disorders, CNS-related disorders, focal
 CC brain disorders, global-diffuse cerebral disorders and other
 CC neurological and cerebrovascular disorders. The CNS disorders include
 CC Alzheimer's disease, senile dementia, Huntington's disease, amyotrophic
 CC lateral sclerosis, Parkinson's, Gilles de la Tourette's syndrome,
 CC autonomic function disorders such as hypertension and sleep disorders,
 CC neuropsychiatric disorders, psychoactive substance use disorders,
 CC anxiety, and bipolar affective disorder.

XX Sequence 390 AA;

Query Match 92.4%; Score 2076; DB 22; Length 390;
 Best Local Similarity 100.0%; Pred. No. 1.6e-202;

Matches 390; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 QY 34 VHMPTKAVDPEAFNFISEIIHQHGYPCPEYEVATEDGYILSVNRI PRGLVQPKKTSRPV 93
 Db 1 VHMPTKAVDPEAFNFISEIIHQHGYPCPEYEVATEDGYILSVNRI PRGLVQPKKTSRPV 60
 QY 94 VLLQHLVGGASNIWISNLPNNISLGFILADAGFDVWNGSRGNASRKHKTLSIDQEFWA 153
 Db 61 VLLQHLVGGASNIWISNLPNNISLGFILADAGFDVWNGSRGNASRKHKTLSIDQEFWA 120
 QY 154 FSYDEMAREFDPVAVINFILQKTGOEKIYVYVGSQGTMTGFIAPSTMPPELAQIKMYFALA 213
 Db 121 FSYDEMAREFDPVAVINFILQKTGOEKIYVYVGSQGTMTGFIAPSTMPPELAQIKMYFALA 180
 QY 214 PIATVYKIAKSPGTFKLLPDMTKGLFGKCFYQTRFLRQVLYICGVILDOICSNIM 273
 Db 181 PIATVYKIAKSPGTFKLLPDMTKGLFGKCFYQTRFLRQVLYICGVILDOICSNIM 240
 QY 274 LLLGGFTNNMNSRASVYAAHTLAGTSVQNIHLWQAVNSGELRAFDMGSETKXLEKCN 333
 Db 241 LLLGGFTNNMNSRASVYAAHTLAGTSVQNIHLWQAVNSGELRAFDMGSETKXLEKCN 300
 QY 334 QPTEVRYRVRDMTPTAMWTGGQDWSNPEDVKLLSEVTNLIYHKNIPWAHVDFTWGL 393
 Db 301 QPTEVRYRVRDMTPTAMWTGGQDWSNPEDVKLLSEVTNLIYHKNIPWAHVDFTWGL 360
 QY 394 DAPHRMYNEIHLMOQEBTNLSQRCBAVL 423
 Db 361 DAPHRMYNEIHLMOQEBTNLSQRCBAVL 390
 RESULT 5
 ID ABG31894 standard; Protein; 423 AA.
 XX
 AC ABG31894;
 XX
 DT 05-NOV-2002 (first entry)
 XX
 DE Human lipid-associated molecule, LIPAML.
 XX
 KW Human; lipid-associated molecule; LIPAM; cancer; leukaemia;
 KW cardiovascular disorder; atherosclerosis; hypertension; stroke;
 KW angina pectoris; neurological disorder; Alzheimer's disease; epilepsy;
 KW dementia; Parkinson's disease; multiple sclerosis; anxiety; psoriasis;
 KW autoimmune disorder; inflammatory disorder; anaemia; asthma; trauma;
 KW acquired immunodeficiency syndrome; AIDS; bronchitis; diabetes mellitus;
 KW emphysema; Crohn's disease; atopic dermatitis; rheumatoid arthritis;
 KW gastrointestinal disorder; lipid metabolism disorder; transgenic animal.
 XX
 OS Homo sapiens.
 XX
 PN WO200246418-A2.
 XX
 PD 13-JUN-2002.
 XX
 PF 04-DEC-2001; 2001WO-US47430.
 XX
 PR 08-DEC-2000; 2000US-254505P.
 PR 15-DEC-2000; 2000US-256187P.
 PR 22-DEC-2000; 2000US-257908P.
 PR 26-JAN-2001; 2001US-264429P.
 XX
 KW (INCY-) INCYTE GENOMICS INC.
 XX
 PI Griffin JA, Gandhi AR, Ramkumar J, Tang YT, Ding L, Yue H;
 PI Gietzen KJ, Sapperstein SK, Honchell CD, Bruns CM, Duggan BM;
 PI Xu Y, Lee S;
 XX
 DR WPI; 2002-619022/66.
 DR N-PSDB; ABR90871.
 XX
 PT Novel polypeptide of human lipid associated molecule, useful for

PT diagnosing, treating and preventing cancer, hypercholesterolaemia,
PT cirrhosis, myocardial infarction, Parkinson's disease, asthma,
PT psoriasis, gastritis -

PS Claim 1; Page 111-112; 126pp; English. xx

The invention relates to an isolated polypeptide (I) of human lipid-associated molecule (LIPAM), and an isolated polynucleotide (II) encoding (I). (I), (II) and antibody to (II) are useful for treating a disease or condition associated with decreased or increased expression of functional LIPAM. (I) or (II) is useful for diagnosing, treating or preventing cancer e.g. leukemia, cancer of adrenal gland, bladder, bone, bone marrow, brain, ovary, etc.; cardiovascular disorders e.g. atherosclerosis, hypertension, Raynaud's disease, angina pectoris, myocardial infarction, rheumatic fever, cardiomyopathy, pericarditis, chronic obstructive pulmonary disease, etc.; neurological disorders e.g. Alzheimer's disease, stroke, epilepsy, dementia, Parkinson's disease, multiple sclerosis, Creutzfeldt-Jakob disease, myasthenia gravis, anxiety, diabetic neuropathy, etc.; autoimmune/inflammatory disorders e.g. acquired immunodeficiency syndrome (AIDS), Addison's disease, anaemia, asthma, bronchitis, diabetes mellitus, emphysema, Crohn's disease, atopic dermatitis, rheumatoid arthritis, psoriasis, systemic lupus erythematosus, trauma, etc.; and gastrointestinal disorders e.g. gastritis, anorexia, nausea gastroenteritis, ulcerative colitis, cholecystitis, hepatitis, cirrhosis, hepatoma, diarrhoea, constipation, etc. and disorders of lipid metabolism e.g., fatty liver, cholestasis, Fabry's disease, hypercholesterolaemia. (I) or (II) is useful for assessing the effects of exogenous compounds on the expression of (I) or (II). (I) is useful in number of drug screening techniques, and to analyse the proteome of a tissue or cell type. (I) or antibody to (II) is useful as elements on a microarray. (II) is useful for creating knockin humanized animals or transgenic animals to model human diseases, in somatic or germline gene therapy, to generate a transcript image of a tissue or cell type, for detecting differences in the chromosomal location due to translocation, inversion, etc., among normal, carrier or affected individuals, and as hybridisation probes for mapping naturally occurring genomic sequences. The present sequence represents the amino acid sequence of a human lipid-associated molecule (LIPAM) of the invention.

AA	Sequence	423 AA;
SQ		

Query Match	89.6%;	Score	2013.5;	DB	23;	Length	423;
Best Local Similarity	92.1%;	Pred.	No.	4.2e-136;			
Matches	384;	Conservative	9;	Mismatches	13;	Indels	11;
							Gaps
QY	17	MYLLIIVAMFQBNVNS	----	VHMPKAVDP	PEAFNISEII	IQHOGYPC	EESEVATEDGYIL 73
DB	8	MYLLITTC	LI	CGTNAGG	FDLENE	-VNP	EVMMNTSEII
QY	74	SVNRI	PRGLVQPKT	GS	RPVVLL	QHLGVGG	ASNWISNL
DB	67	SVNRI	PRGLVQPKT	GS	RPVVLL	QHLGVGG	ASNWISNL
QY	134	GNWSE	RKHTLSI	DD	DEFWAF	SYDEMAR	FDLP
DB	127	GNWSE	RKHTLSI	DD	DEFWAF	SYDEMAR	FDLP
QY	194	IAFST	MPPELA	QKI	KMYF	ALAPIAT	IVYGHAKS
DB	187	IAFST	MPPELA	QKI	KMYF	ALAPIAT	IVYGHAKS
QY	254	QLVYI	LCGV	IL	DQI	CSNIM	LLGFGFNTNNNM
DB	247	QLVYI	LCGV	IL	DQI	CSNIM	LLGFGFNTNNNM
QY	307	HWSQ	VNSGEL	RA	FD	WGS	ETK
DB	307	HWSQ	VNSGEL	RA	FD	WGS	ETK
QY	367	MLLS	EVNTNLI	YHKNI	PE	WAH	VDFT
DB	367	MLLS	EVNTNLI	YHKNI	PE	WAH	VDFT

PT New isolated human lipase peptides and encoding nucleic acids, useful
PT for diagnosing and treating disorders mediated by human lipase
PT proteins, e.g. cancer of the colon, kidneys, skin and testis, and
PT Burkitt's lymphoma -
XX
PS Claim 1; Fig 2; 79pp; English.
XX
CC The present invention relates to a new lipase peptide. The methods and
CC compositions of the present invention are useful for diagnosing and
CC treating disorders mediated by the human lipase protein, such as cancer
CC of the colon, kidneys, skin, brain and testis, and Burkitt's lymphoma.
CC The present amino acid sequence represents the human lipase protein of
CC the invention. This sequence is encoded by the human lipase gene
CC located on chromosome 10.
XX
SQ Sequence 398 AA;

Query Match 88.6%; Score 1991; DB 23; Length 398;
Best Local Similarity 100.0%; Pred. No. 7.6e-194;
Matches 374; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 50 SEIIHQGYPCPEYEVATEDGYILSVNRIIPRGLVQPKTGSFVLLQHLGVLGASNWS 109
DB 25 SEIIHQGYPCPEYEVATEDGYILSVNRIIPRGLVQPKTGSFVLLQHLGVLGASNWS 84
QY 110 NLPNNSLGFILADAGFDVWNGSRGNWSRKHKTLSIDQDEFWAFSYDEMARDLPAVIN 169
DB 85 NLPNNSLGFILADAGFDVWNGSRGNWSRKHKTLSIDQDEFWAFSYDEMARDLPAVIN 144
QY 170 FILQKTQGEKIYYVYSGQTTMGFIAPSTMPELAQIKMYFALAPIATVKHAKSPGTKFL 229
DB 145 FILQKTQGEKIYYVYSGQTTMGFIAPSTMPELAQIKMYFALAPIATVKHAKSPGTKFL 204
QY 230 LLPDMNKGFLGKKEFLYQTRFLRQLVILCGOVIDQICSNIMLLGGFTNNMNSRA 289
DB 205 LLPDMNKGFLGKKEFLYQTRFLRQLVILCGOVIDQICSNIMLLGGFTNNMNSRA 264
QY 290 SVYAAHTLAGTSVQNIILHWSQAVNSGELRAFDMGSETKLEKNCQPTPVRYRDMVTPT 349
DB 265 SVYAAHTLAGTSVQNIILHWSQAVNSGELRAFDMGSETKLEKNCQPTPVRYRDMVTPT 324
QY 350 AMWTGGQDWLSNPEDVKMLLSEVNTLIYHKNIPEWAHVDFIWLGLDAPHRMYNEIHLMOQ 409
DB 325 AMWTGGQDWLSNPEDVKMLLSEVNTLIYHKNIPEWAHVDFIWLGLDAPHRMYNEIHLMOQ 384
QY 410 EETNLSQGRCEAVL 423
DB 385 EETNLSQGRCEAVL 398

RESULT 7
AAU99164
ID AAU99164 standard; Protein; 398 AA.
XX
XX AAU99164;
AC AAU99164;
XX
DT 24-SEP-2002 (first entry)
XX
DE Human lysosomal acid lipase #1.
XX
KW Human; enzyme; lysosomal acid lipase; lipid malabsorption illness;
KW cystic fibrosis; alcoholism; heart disease; heart attack;
KW Wolman disease; cholesterol ester storage disease; brain injury;
KW mood disorder; anxiety disorder; thought disorder; volition disorder;
KW obesity; cancer; neurogenic disorder; myopathic disorder; COPD;
KW cardiovascular disorder; chronic obstructive pulmonary disease; diabetes;
KW anorexia; osteoarthritis; Alzheimer's disease; Parkinson's disease;
KW peripheral nervous system disorder.
XX
OS Homo sapiens.
XX
FN WO200236731-A2.

XX 10-MAY-2002.
XX 30-OCT-2001; 2001WO-EPI2518.
XX 31-OCT-2000; 2000US-244170P.
PR 29-MAY-2001; 2001US-293516P.
XX (FARB) BAYER AG.
XX Xiao Y;
XX WPI; 2002-519248/55.
DR N-PSDB; ABK86569.
XX
PT Novel human lysosomal acid lipase polypeptide, useful for treating
PT cancer, diabetes, obesity, chronic obstructive pulmonary disease,
PT peripheral or central nervous system disorder or cardiovascular
PT disorder -
XX
PS Claim 25; Fig 2; 126pp; English.

XX The invention relates to a purified human lysosomal acid lipase
CC polypeptide. Also included are the polynucleotide encoding the
CC lipase (or its fragment, derivative, allele or sequence at least 60%
CC identical to it), vectors, host cells, a reagent (e.g. an antisense
CC oligonucleotide), which binds to the lipase or polynucleotide (used for
CC detection and modulating/reducing the lipase activity) and an anti-lipase
CC antibody. The lipase and polynucleotide are useful for identifying
CC therapeutic agents that either increase or decrease the lipase activity.
CC The identified agent, the lipase and polynucleotide are useful for
CC treatment of a disease such as lipid malabsorption illness,
CC cystic fibrosis, alcoholism, heart disease, heart attack,
CC Wolman disease, cholesterol ester storage disease, brain injury,
CC mood disorder, anxiety disorder, thought disorder, volition disorder,
CC sleep disorder, neurogenic disorder, myopathic disorder, obesity, cancer,
CC chronic obstructive pulmonary disease (COPD), diabetes,
CC cardiovascular disorder, Alzheimer's disease, Parkinson's disease,
CC anorexia, osteoarthritis, a central nervous system disorder and
CC a peripheral nervous system disorder. The present sequence is the
CC human lysosomal lipase #1.
XX
SQ Sequence 398 AA;

Query Match 88.6%; Score 1991; DB 23; Length 398;
Best Local Similarity 100.0%; Pred. No. 7.6e-194;
Matches 374; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 50 SEIIHQGYPCPEYEVATEDGYILSVNRIIPRGLVQPKTGSFVLLQHLGVLGASNWS 109
DB 25 SEIIHQGYPCPEYEVATEDGYILSVNRIIPRGLVQPKTGSFVLLQHLGVLGASNWS 84
QY 110 NLPNNSLGFILADAGFDVWNGSRGNWSRKHKTLSIDQDEFWAFSYDEMARDLPAVIN 169
DB 85 NLPNNSLGFILADAGFDVWNGSRGNWSRKHKTLSIDQDEFWAFSYDEMARDLPAVIN 144
QY 170 FILQKTQGEKIYYVYSGQTTMGFIAPSTMPELAQIKMYFALAPIATVKHAKSPGTKFL 229
DB 145 FILQKTQGEKIYYVYSGQTTMGFIAPSTMPELAQIKMYFALAPIATVKHAKSPGTKFL 204
QY 230 LLPDMNKGFLGKKEFLYQTRFLRQLVILCGOVIDQICSNIMLLGGFTNNMNSRA 289
DB 205 LLPDMNKGFLGKKEFLYQTRFLRQLVILCGOVIDQICSNIMLLGGFTNNMNSRA 264
QY 290 SVYAAHTLAGTSVQNIILHWSQAVNSGELRAFDMGSETKLEKNCQPTPVRYRDMVTPT 349
DB 265 SVYAAHTLAGTSVQNIILHWSQAVNSGELRAFDMGSETKLEKNCQPTPVRYRDMVTPT 324
QY 350 AMWTGGQDWLSNPEDVKMLLSEVNTLIYHKNIPEWAHVDFIWLGLDAPHRMYNEIHLMOQ 409
DB 325 AMWTGGQDWLSNPEDVKMLLSEVNTLIYHKNIPEWAHVDFIWLGLDAPHRMYNEIHLMOQ 384
QY 410 EETNLSQGRCEAVL 423

Query Match 68.4%; Score 1536.5; DB 23; Length 371;
Best Local Similarity 80.9%; Pred. No. 1.4e-147;

CC hypertension.
XX
SQ Sequence 399 AA;

Query Match 57.4%; Score 1289; DB 22; Length 399;
Best Local Similarity 63.5%; Pred. No. 2.6e-122;
Matches 235; Conservative 55; Mismatches 80; Indels 0; Gaps 0;

QY 40 AVDPEAFNMISIIHQGYPCPEYEVATEDGYILSVNRIPLRGVLPKKTGSRPVVLLQHG 99
DB 28 AVDPETNMVSEIIISYWGPFSEYLVETEDGYILCLNRIPLRGKNSDKGPKPVVFLQHG 87

QY 100 LVGASNWTLSNPNNSLGLADAGFDVWNGSRGNWSRKHKTLSIDQDEFWAFSYDEM 159
DB 88 LLADSSNWTNLANSLSGLFADAGFDVWNGSRGNWTSRKHKTLSVQDEFWAFSYDEM 147

QY 160 ARFDLPVAVINFILOKTQGBKIIYVYVGSQGTMGFIAPSTMPPELAQIKMYFALAPIATVK 219
DB 148 AKYDLPASINFILOKTQGBKIIYVYVGSQGTMGFIAPSTMPPELAQIKMYFALAPIATVK 207

QY 220 HAKSPGTEKLLPDMMIKGLFGKKEFLYQTRFLQVILYLCQVILDOICSNIMLLGGF 279
DB 208 FCTSPMAKLGRLPDHLIKDLFGDKFELPOSALFKMLGTHVTHVILKELCGNLCFLQGF 267

QY 280 NTNNMNSRASVYAAHTLAGTSVQNILHWSQAVNSGELRAFDWGSSETKNLEKCNQPTPVR 339
DB 268 NERNLMSRVVDVYTHSPAGTSVQNMLHWSQAVKQKQAFDWSGSAKNYFHYNQSYPT 327

QY 340 YRVDMTPTAMWTGGQDWSNPEDVKMLLSEVTNLIYHKNIPENAHVDFIWLGDAPHRM 399
DB 328 YNVKDMLVPTAVWSGGHDLADVDVYNILLTQITNLVPHESIPWEHLDPIWGLDAPWRL 387

QY 400 YNEIHLMOQ 409
DB 388 YNKIINLMEK 397

RESULT 11
AA018227
ID AA018227 standard; Protein; 399 AA.
AC AAB66061;
XX
DT 30-MAR-2001 (first entry)
XX
DE Human lysosomal acid lipase protein.
XX
KW Gene therapy; TANGO protein; INTERCEPT protein; neurological disorder;
KW central nervous system; focal brain disorder; bipolar affective disorder;
KW global-diffuse cerebral disorder; cerebrovascular; Alzheimer's disease;
KW senile dementia; Huntington's disease; amyotrophic lateral sclerosis;
KW Parkinson's; Gilles de la Tourette's syndrome; hypertension; sleep;
KW neuropsychiatric; psychoactive substance use; anxiety.
XX
OS Homo sapiens.
XX
XX WO200077239-A2.
XX
XX 21-DEC-2000.
XX
XX 24-MAY-2000; 2000WO-US14858.
XX
XX 14-JUN-1999; 99US-0333159.
XX
XX (MILL-) MILLENNIUM PHARM INC.
XX
XX McCarthy SA, Fraser CC, Sharp JD, Barnes TM;
XX
XX WPI; 2001-032313/04.
XX
XX TANGO and INTERCEPT nucleic acids, proteins, and antibodies, useful for
PT screening assays and diagnostic assays and for the treatment of
PT neurological diseases such as Alzheimer's, Parkinson's and Huntington's

PT disease -
XX Claim 8; Fig 6; 359pp; English.
XX
CC The present invention relates to TANGO or INTERCEPT proteins and coding
CC sequences (see AAF45121-F45136 and AAF45138-F45139 and AAB66031-B66057;
CC AAB66064-B66083 and AAB66085). The TANGO/INTERCEPT proteins and coding
CC sequences are useful for the treatment of neurological disorders such as
CC central nervous system (CNS) disorders, CNS-related disorders, focal
CC brain disorders, global-diffuse cerebral disorders and other
CC neurological and cerebrovascular disorders. The CNS disorders include
CC Alzheimer's disease, senile dementia, Huntington's disease, amyotrophic
CC lateral sclerosis, Parkinson's, Gilles de la Tourette's syndrome,
CC autonomic function disorders such as hypertension and sleep disorders,
CC neuropsychiatric disorders, psychoactive substance use disorders,
CC anxiety, and bipolar affective disorder. The present sequence is a
CC sequence used in a sequence homology alignment with the TANGO/INTERCEPT
CC sequences of the present invention.
XX
SQ Sequence 399 AA;

Query Match 57.4%; Score 1289; DB 22; Length 399;
Best Local Similarity 63.5%; Pred. No. 2.6e-122;
Matches 235; Conservative 55; Mismatches 80; Indels 0; Gaps 0;

QY 40 AVDPEAFNMISIIHQGYPCPEYEVATEDGYILSVNRIPLRGVLPKKTGSRPVVLLQHG 99
DB 28 AVDPETNMVSEIIISYWGPFSEYLVETEDGYILCLNRIPLRGKNSDKGPKPVVFLQHG 87

QY 100 LVGASNWTLSNPNNSLGLADAGFDVWNGSRGNWSRKHKTLSIDQDEFWAFSYDEM 159
DB 88 LLADSSNWTNLANSLSGLFADAGFDVWNGSRGNWTSRKHKTLSVQDEFWAFSYDEM 147

QY 160 ARFDLPVAVINFILOKTQGBKIIYVYVGSQGTMGFIAPSTMPPELAQIKMYFALAPIATVK 219
DB 148 AKYDLPASINFILOKTQGBKIIYVYVGSQGTMGFIAPSTMPPELAQIKMYFALAPIATVK 207

QY 220 HAKSPGTEKLLPDMMIKGLFGKKEFLYQTRFLQVILYLCQVILDOICSNIMLLGGF 279
DB 208 FCTSPMAKLGRLPDHLIKDLFGDKFELPOSALFKMLGTHVTHVILKELCGNLCFLQGF 267

QY 280 NTNNMNSRASVYAAHTLAGTSVQNILHWSQAVNSGELRAFDWGSSETKNLEKCNQPTPVR 339
DB 268 NERNLMSRVVDVYTHSPAGTSVQNMLHWSQAVKQKQAFDWSGSAKNYFHYNQSYPT 327

QY 340 YRVDMTPTAMWTGGQDWSNPEDVKMLLSEVTNLIYHKNIPENAHVDFIWLGDAPHRM 399
DB 328 YNVKDMLVPTAVWSGGHDLADVDVYNILLTQITNLVPHESIPWEHLDPIWGLDAPWRL 387

QY 400 YNEIHLMOQ 409
DB 388 YNKIINLMEK 397

RESULT 11
AA018227
ID AA018227 standard; Protein; 399 AA.
AC AAB66061;
XX
DT 30-MAR-2001 (first entry)
XX
DE Human lysosomal acid lipase protein.
XX
KW Gene therapy; TANGO protein; INTERCEPT protein; neurological disorder;
KW central nervous system; focal brain disorder; bipolar affective disorder;
KW global-diffuse cerebral disorder; cerebrovascular; Alzheimer's disease;
KW senile dementia; Huntington's disease; amyotrophic lateral sclerosis;
KW Parkinson's; Gilles de la Tourette's syndrome; hypertension; sleep;
KW neuropsychiatric; psychoactive substance use; anxiety.
XX
OS Homo sapiens.
XX
XX WO200077239-A2.
XX
XX 21-DEC-2000.
XX
XX 24-MAY-2000; 2000WO-US14858.
XX
XX 14-JUN-1999; 99US-0333159.
XX
XX (MILL-) MILLENNIUM PHARM INC.
XX
XX McCarthy SA, Fraser CC, Sharp JD, Barnes TM;
XX
XX WPI; 2001-032313/04.
XX
XX TANGO and INTERCEPT nucleic acids, proteins, and antibodies, useful for
PT screening assays and diagnostic assays and for the treatment of
PT neurological diseases such as Alzheimer's, Parkinson's and Huntington's

PF 29-MAR-2001; 2001US-0820001.
 XX 29-MAR-2001; 2001US-0820001.
 XX (PEKE) PE CORP NY.
 XX Merkulo GV, Ketchum KA, Di Francesco V, Beasley EM;
 XX WPI; 2002-478445/51.
 XX Isolated nucleic acids encoding a lysosomal acid lipase protein useful
 PT for the prevention, diagnosis and treatment of severe infantile-onset
 PT Wolman disease and late-onset cholesteryl ester storage disease -
 XX acid lipase (LIPA) gene. The present sequence is a protein shown in the
 XX exemplification of the invention.
 XX Sequence 392 AA;
 CC
 CC The present invention provides the protein, coding and genomic sequences
 CC of a human lysosomal acid lipase. The sequences can be used in the
 CC identification of modulators of lipase activity in cells and tissues that
 CC express the lipase, particularly the severe infantile-onset Wolman
 CC disease and the milder late-onset cholesteryl ester storage disease
 CC (CSD), which are caused by mutations in different parts of the lysosomal
 CC acid lipase (LIPA) gene. The present sequence is a protein shown in the
 CC exemplification of the invention.
 XX
 XX Sequence 392 AA;
 CC
 CC Query Match 57.1%; Score 1282; DB 23; Length 392;
 CC Best Local Similarity 63.8%; Pred. No. 1.3e-121;
 CC Matches 234; Conservative 53; Mismatches 80; Indels 0; Gaps 0;
 QY 40 AVDPFAFNISEIIHQGYPCCEVEVATEGYYLLSVNRIPRGLVQPKTGRVNVLLQHG 99
 DB 26 AVDPETNNVSEIIISYWGFPSEYLVETEDGYYLLSVNRIPRGLVQPKTGRVNVLLQHG 85
 QY 100 LVGGASNWSNLNPNNSLGFILADAGFDVWNGNSRGNWSRKHKTLSIDQDEFWAFSDYM 159
 DB 86 LLADSSNWNVTNLSLGFILADAGFDVWNGNSRGNWSRKHKTLSVSDQDEFWAFSDYM 145
 QY 160 ARPDLPAVINFIQTKQEKIYYGYSGQTMGFIAFTWPELAQKIKNYFALAPATVK 219
 DB 146 AKYDLPASINFIQTKQEKIYYGYSGQTMGFIAFTWPELAQKIKNYFALAPATVK 205
 QY 220 HAKSPGKFKLLPDMWIKGLGKKEFLYQTRFLRQLVYLCGOVILQICSNIMLLGGF 279
 DB 206 FCTSPMAKLGRLPHLILKDLGDEKFLPQSAFLKWLGTHTVCTVILKELCNLCFLCGF 255
 QY 280 NTNNNSRASVYAAHTLAGTSVONILHWSQAVNSGELRPFADMGSETKNEKCNQPTPVR 339
 DB 266 NERNLANSRVDDVYTHSPAGTSVQNMHLHWSQAVKFKQFQAFDWMGSSAKNYFHYNQSPPT 325
 QY 340 YRVDMTVPMTAMWTGGODLSNPEDVKMLLSEVTNLIYHKNIPWAHVDPIWGLDAPHRM 399
 DB 326 YNVKMLVPTAVNSGGHDLADVYDVNILLTQITNLVNFHESIPWEHLDFIWGLDAPWRL 385
 QY 400 YNEIHL 406
 DB 386 YNKIINL 392
 RESULT 12
 AAU30498
 ID AAU30498 standard; Protein; 449 AA.
 XX
 XX AAU30498;
 AC
 XX
 DT 18-DEC-2001 (first entry)
 XX
 XX Novel human secreted protein #989.
 DE Human; vaccination; gene therapy; nutritional supplement;
 KW stem cell proliferation; haematopoiesis; nerve tissue regeneration;
 KW immune suppression; immune stimulation; anti-inflammatory; leukaemia.

XX Homo sapiens.
 OS WO200179449-A2.
 XX
 XX 25-OCT-2001.
 PD
 XX
 XX 16-APR-2001; 2001WO-US08656.
 PF
 XX
 XX 18-APR-2000; 2000US-0552929.
 PR 26-JAN-2001; 2001US-0770160.
 XX (HYSE-) HYSEQ INC.
 PA
 XX
 XX Tang YT, Liu C, Drmanac RT;
 PI WPI; 2001-611725/70.
 XX
 XX Nucleic acids encoding a range of human polypeptides, useful in genetic
 PT vaccination, testing and therapy -
 XX
 XX Claim 20; Page 301; 765pp; English.
 XX
 XX The invention relates to novel human secreted polypeptides. The
 CC polypeptides and antibodies to the polypeptides are useful for
 CC determining the presence of or predisposition to a disease associated
 CC with altered levels of polypeptide. The polypeptides are also useful for
 CC identifying agents (agonists and antagonists) that bind to them. Cells
 CC expressing the proteins are useful for identifying a therapeutic agent
 CC for use in treatment of a pathology related to aberrant expression or
 CC physiological interactions of the polypeptide. Vectors comprising
 CC the nucleic acids encoding the polypeptides and cells genetically
 CC engineered to express them are also useful for producing the proteins.
 CC The proteins are useful in genetic vaccination, testing and
 CC therapy, and can be used as nutritional supplements. They may be used to
 CC increase stem cell proliferation; to regulate haematopoiesis; and in
 CC bone, cartilage, tendon and/or nerve tissue growth or regeneration;
 CC immune suppression and/or stimulation; as anti-inflammatory agents; and
 CC in treatment of leukaemias. AAU29510-AAU33304 represent the amino acid
 CC sequences of novel human secreted proteins of the invention.
 XX
 XX Sequence 449 AA;
 SQ
 Query Match 53.0%; Score 1191.5; DB 22; Length 449;
 Best Local Similarity 57.4%; Pred. No. 2.7e-112;
 Matches 230; Conservative 59; Mismatches 103; Indels 3; Gaps 2;
 QY 12 SHRMEWMLLILVAYNFQNRVNSVHPTK--AVDPEAFNISEIIHQGYPCCEVEVATED 69
 DB 47 SSRMKRVLGLVVCVLTSLHSEGSGGKLTAVDPETNNVSEIIISYWGFPSEYLVETED 106
 QY 70 GYILSVNRIPLGLVQPKTGRVNVLLQHGVLGGASNWSNLNPNNSLGFILADAGFDVW 129
 DB 107 GYILCNLIPGRKNHSDKGPVVFLOHGLADSSNWNVTNLSRGLFILADAAIDVW 166
 QY 130 GNSRGNWSRKHKTLSIDQDEFWAFSDYDEMARFDPVAVINFILQTKQEKIYYGYSGQT 189
 DB 167 GNTRGNTWSPKHKTLSVSDQDEFWAFSDYDEMAKYDLPASINFLNKTGQEVYVYVGHSGQT 226
 QY 190 TMGPTAFSTMPBELAQIKNYFALAPATVYHAKSPGKFKLLPDMWIKGLGKKEFLYQTR 249
 DB 227 TIGTFAFSQMLELAKGLXKFFAWGPFVAFCTSPMAKLGRLPHLILKDLGDEKFLPQ 286
 QY 250 RFLRQL-VYILCGOVILQICSNIMLLIGFNTNNMSRASVYAAHTLAGTSVQNIILH 308
 DB 287 AFWKAVGYPHLATHVILKELCNLCFLCGFNERNLANSRVDDVYTHSPAGTFFQNKXHW 346
 QY 309 SOAVNSGELRPFADMGSETKNEKCNQPTPVRVYRDMTVPMTAMWTGGODLSNPEDVKML 368
 DB 347 SOAVKFKQFQAFDWMGSSAKNYFHYNQSPPTVNVKMLVPTAVYTGTHDMLDYVGVNIX 406
 QY 369 LSEVTNLIYHKNIPWAHVDPIWGLDAPHRMNIILHMQ 409
 XX
 XX

Db 407 LQITWLVFHSIPWEHLDFIWLGLDAPWRLNKLINLRK 447

RESULT 13

AAR37302
ID AAR37302 standard; Protein; 398 AA.

XX AAR37302;

XX 25-MAR-2003 (updated)

DT 20-SEP-1993 (first entry)

DE RGL precursor.

XX Rabbit gastric lipase; RGL; prGLN2.1; fat; bioconversion;
KW hydrolysis; transesterification.

OS Oryctolagus cuniculus.

XX Key Location/Qualifiers
FT Protein 23..398

FT /note= "claim 1; page 10-11"

PN EP542629-A1.

XX 19-MAY-1993.

XX 12-NOV-1992; 92EP-0403055.

XX 13-NOV-1991; 91FR-0013948.

XX (LJOU) INST RECH JOUVEINAL.

XX Benicourt C, Blanchard C, Junien J;

XX WPI; 1993-161080/20.

DR N-PSDB; AAQ42310.

XX Rabbit gastric lipase, its precursor and their DNA - useful for
PT treating conditions linked to gastric lipase deficiency, such as
PT mucoviscidiosis and pancreatic exocrine insufficiency

XX Claim 1; Fig 7; 31pp; French.

XX RGL, opt. used with other lipases, are useful therapeutically
CC (1) to facilitate absorption of ingested fats in patients deficient
CC in endogenous gastric lipase and (2) to treat disorders caused by
CC inadequate lipase prodn. esp. mucoviscidiosis or pancreatic exocrine
CC deficiency. It can also be used (partic. when immobilised) for
CC enzymatic bioconversion, e.g. hydrolysis or transesterification.
CC (Updated on 25-MAR-2003 to correct PN field.)

XX Sequence 398 AA;

Query Match 51.9%; Score 1166; DB 14; Length 398;
Best Local Similarity 54.6%; Pred. No. 8.8e-110;
Matches 219; Conservative 70; Mismatches 102; Indels 10; Gaps 3;

QY 17 MWLLLVAYMFORNVSVH-----MPTKAVDPEAFNNISEIIOHOGYCEEEVATEDG 70
DB 1 MWVLFMVAALISA-LGTHGLFGKSAFT---NPEVNNISQMSYWGYPSEKYEVTEDG 56

QY 71 YILSVNRIPRGLVQPKTKTSRPPVLLQHGLVGGASNWSLNPNSLGFILADAGFDVWVG 130
DB 57 YILEVNRIPYKKGKSGNRGQRPVVFLOHGLLASASNWSLNPNSLAFILADAGYGVWLG 116

QY 131 NSRGNASRKHKLISIQDDEFAFSDVEMARPLPAVINFILOKTQEKLYYVGSQGT 190
DB 117 NSRGNWNRNLYYSPDSVFFAFSFDENAKYDLPATIDFIVKETQCKLHYVGHSGQGT 176

QY 191 MGFAFSTMPELAKIKWYFALAPIATVGHAKSPGTFKLLLPDMWIKGLFGKKEFLYQTR 250
DB 177 IGFAFSTNPKLABRIKTFALAPVAIVKTKSLVNNKLRFPPTMFKIFGDKIFYPHNF 236

QY 251 FLRLVYLGQVILDOICSNIMLLGGFTNNMNSRASVVAATLACTSVQNILHWSQ 310
DB 237 FDFLATQVCSRETLNVCISNALFIICGFDNSANLNSRLDVTYVSNPACTSVQNMLHWIQ 296
QY 311 AVNSGELRAPDWGSETKNLEKNQPPRYRVRDMTVPTAMMTGGODWLSNEDVVKLLUS 370
DB 297 AVLSGNFQAFNMGSPAQNVVHFNQTPPYNYNTAMNVPYAVWSGGNDWLDADQDVDDLLP 356
QY 371 EVTNLYHKNIPEWAHVDFIWLGLDAPHRMYNEIHLMQOEE 411
DB 357 KLSNLYHKELIPYNHLDFFIWNWAPQEVNIELISMAKDK 397

RESULT 14

AAB66068

ID AAB66068 standard; Protein; 221 AA.

XX AAB66068;

XX 30-MAR-2001 (first entry)

DE Human TANGO 294 extracellular domain.

XX TANGO protein; INTERCEPT protein; neurological disorder;

KW central nervous system; focal brain disorder; bipolar affective disorder;
KW global-diffuse cerebral disorder; cerebrovascular; Alzheimer's disease;
KW senile dementia; Huntington's disease; amyotrophic lateral sclerosis;
KW Parkinson's; Gilles de la Tourette's syndrome; hypertension; sleep;
KW neuropsychiatric; psychoactive substance use; anxiety.

OS Homo sapiens.

XX WO2000077239-A2.

XX 21-DEC-2000.

XX 24-MAY-2000; 2000WO-US14858.

XX 14-JUN-1999; 99US-0333159.

XX (MILL-) MILLENNIUM PHARM INC.

XX McCarthy SA, Fraser CC, Sharp JD, Barnes TM;

XX WPI; 2001-032313/04.

XX TANGO and INTERCEPT nucleic acids, proteins, and antibodies, useful for
PT screening assays and diagnostic assays and for the treatment of
PT neurological diseases such as Alzheimer's, Parkinson's and Huntington's
PT disease -

XX Claim 8; Page 326; 359pp; English.

XX The present invention relates to TANGO or INTERCEPT proteins and coding
CC sequences (see AAF45121-F45136 and AAF45138-F45139 and AAB66031-B66057,
CC AAB66054-B66083 and AAB66085). The TANGO/INTERCEPT proteins and coding
CC sequences are useful for the treatment of neurological disorders such as
CC central nervous system (CNS) disorders, CNS-related disorders, focal
CC brain disorders, global-diffuse cerebral disorders and other
CC neurological and cerebrovascular disorders. The CNS disorders include
CC Alzheimer's disease, senile dementia, Huntington's disease, amyotrophic
CC lateral sclerosis, Parkinson's, Gilles de la Tourette's syndrome,
CC autonomic function disorders such as hypertension and sleep disorders,
CC neuropsychiatric disorders, psychoactive substance use disorders,
CC anxiety, and bipolar affective disorder.

XX Sequence 221 AA;

Query Match 51.8%; Score 1165; DB 22; Length 221;

Best Local Similarity 100.0%; Pred. No. 4.5e-110;

Matches 221; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 34 VHMPTKAVDPEAFMNISEIIHQGYPCPEYEVATEDGYILSVNRIPRGLVQPKTGSRPV 93
 DB 1 VHMPTKAVDPEAFMNISEIIHQGYPCPEYEVATEDGYILSVNRIPRGLVQPKTGSRPV 60
 QY 94 VILQGLVGGASWISNLPNNISLGFILADAGFDVWGNISRGNAWKRKHKTLSIDODEFWA 153
 DB 61 VILQGLVGGASWISNLPNNISLGFILADAGFDVWGNISRGNAWKRKHKTLSIDODEFWA 120
 QY 154 FSYDEMARFDLPAVINFILOKTQGEKIYVGVYSGQTTMGFIAPFSTMPPELAQIKMYFALA 213
 DB 121 FSYDEMARFDLPAVINFILOKTQGEKIYVGVYSGQTTMGFIAPFSTMPPELAQIKMYFALA 180
 QY 214 PIATVTHAKSPGKTFLLPDMIMKGLFGKKEFLYOTRFLRQ 254
 DB 181 PIATVTHAKSPGKTFLLPDMIMKGLFGKKEFLYOTRFLRQ 221

RESULT 15

AAFP50322 standard; protein; 395 AA.

XX AC AAP50322;
 XX AC AAP50322;
 DT 25-MAR-2003 (updated)
 DT 17-JAN-1992 (first entry)
 XX DT
 DE Rat lingual lipase protein.
 XX
 KW Lingual lipase; enzyme; EC-3.1.1.3; ss.
 OS Rattus rattus.
 XX OS
 XX OS
 PN GB2142337-A.
 XX
 PD 16-JAN-1985.
 XX
 XX 29-JUN-1984; 84GB-0016581.
 XX
 XX 29-JUN-1984; 84GB-0016581.
 XX
 PA (CLLT) CELLTech LTD.
 PA (WILL/) WILLIAMSON R.
 XX
 PI Carey NH, Williamson R;
 XX
 DR WPI; 1985-014450/03.
 DR N-PSDB; AAN50385.
 XX
 PT New lingual lipase protein for treatment of lipase deficiency - also
 PT new pre-lingual lipase protein and related products
 XX
 PS Disclosure; Fig 5; 15pp; English.
 XX
 CC This protein may be expressed in a transformant host organism
 CC and may be used for the treatment of lipase deficiency.
 CC (Updated on 25-MAR-2003 to correct PR field.)
 CC (Updated on 25-MAR-2003 to correct PA field.)
 XX
 SQ Sequence 395 AA;

Query Match 51.5%; Score 1156.5; DB 6; Length 395;
 Best Local Similarity 54.7%; Pred. No. 8.1e-109;
 Matches 216; Conservative 69; Mismatches 109; Indels 1; Gaps 1;
 QY 17 MWLLILVAVMFO-RNVNSVHMPKAVDPEAFNISEIIHQGYPCPEYEVATEDGYILSV 75
 DB 1 MWLLILVAVMFO-RNVNSVHMPKAVDPEAFNISEIIHQGYPCPEYEVATEDGYILSV 60
 QY 76 NRIPRGLVQPKTGSRPVLLQGLVGGASWISNLPNNISLGFILADAGFDVWGNISRGNA 135
 DB 61 YRPHGKNNSEIGKPEPVVYLQGLVGGASWISNLPNNISLGFILADAGFDVWGNISRGNA 120
 QY 136 AWRKHKTLSIDODEFWAFSYDEMARFDLPAVINFILOKTQGEKIYVGVYSGQTTMGFI 195

DB 121 TWSKKNVYSPDSVEFWAFSDEWAKYDLPATINFIVOKTQGEKIHYVGHVSQRTTIGFIA 180
 QY 196 FSTMPPELAQIKMYFALAPATVTHAKSPGKTFLLPDMIMKGLFGKKEFLYOTRFLRQ 255
 DB 181 FSTNPTLAKKIKTFYALAPVATVKTQSPILKKISFIPTFLFKLMFGKGMFLPHTYEDDFL 240
 QY 256 VIYLCGOVILDOICSNIMLILGGFNTNNMMSASVYAAHTLAGTSVONILHWSQAVNSG 315
 DB 241 GTEVCSREVLDLCSNTLFIKCFDKNLNVSRPDVYLGHNPAGTSVQDFLHWAQLVRS 300
 QY 316 ELRAFDGSGSTKMLEKCNQPTPVRYVRDNTVPTAMWTGGQDLWLSNPEDVYKMLLSEVTNL 375
 DB 301 KFOAFNMGSPSQNNLHYNQKTPPEYDVSAMTVPVAVVNGGNDILADPQDVAMLLPKLSNL 360
 QY 376 IYHKNIPWAHVDPIWGLDAPHERMYNEIHLMOOE 410
 DB 361 LFKKILAYNHLDFIWANDAPQEVYNEIMISMMAED 395

Search completed: February 19, 2004, 14:20:36
 Job time : 37.8598 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2004 Compugen Ltd.

OM protein - protein search, using sw model

Run on: February 19, 2004, 14:17:49 ; Search time 14.5683 Seconds
(without alignments)
2792.326 Million cell updates/sec

Title: US-10-042-431-47

Perfect score: 2247

Sequence: 1 MLETLRQWIVSHRMENWLL.....IHLMQQETNLSQRCRAVL 423

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 283308 seqs, 96168682 residues

Total number of hits satisfying chosen parameters: 283308

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database :

1: PIR.76.*

2: PIR1.*

3: PIR2.*

4: PIR3.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	1290	57.4	399	2	S41408 lysosomal acid lip
2	1290	57.4	399	2	G01416 lysosomal acid lip
3	1164.5	51.8	395	1	L1RTT triacylglycerol li
4	1139	50.7	397	1	J04017 triacylglycerol li
5	1119	49.8	398	2	S07145 triacylglycerol li
6	752.5	33.5	403	2	T33198 hypothetical prote
7	751	33.4	411	2	T32290 hypothetical prote
8	730	32.5	405	2	T22675 hypothetical prote
9	707.5	31.5	405	2	H88930 protein R1011.14
10	692	30.8	411	2	G89074 protein K04A8.5 [i
11	655.5	29.2	426	2	T20480 hypothetical prote
12	574.5	25.6	559	2	J70949 eeg-specific prote
13	505	22.5	443	2	T39540 triglyceride lipas
14	478.5	21.3	1585	2	T31611 hypothetical prote
15	475	21.1	467	2	T41053 triglyceride lipas
16	445	19.8	344	2	E84526 probable lysosomal
17	422.5	18.8	548	2	S37969 probable triacygl
18	412.5	18.4	460	2	T39443 probable triglycer
19	411.5	18.3	413	2	T43170 probable triacygl
20	294	13.1	431	2	D86318 protein P15H18.6 [
21	285	12.7	538	2	S84842 probable membrane
22	271.5	12.1	573	2	S64754 probable membrane
23	234.5	10.4	509	2	G96766 protein lipase F2P
24	149	6.6	62	2	S59904 lipase, pregastric
25	143	6.4	336	2	F83425 probable esterase/
26	126	5.6	987	2	A64474 hypothetical prote
27	113	5.0	412	2	B72391 conserved hypother
28	112.5	5.0	40	2	S19533 triacylglycerol li
29	107	4.8	281	2	A11912 lysophospholipase

ALIGNMENTS

RESULT 1

S41408

Lysosomal acid lipase (EC 3.1.1.-) / sterol esterase (EC 3.1.1.13) precursor - human
C/Species: Homo sapiens (man)

C/Date: 13-Jan-1995 #sequence_revision 13-Jan-1995 #text_change 18-Jun-1999

C/Accession: S41408; A39315; S47187

R/Amets, D.; Merkel, M.; Eckerskorn, C.; Greten, H.

Eur. J. Biochem. 219, 905-914, 1994

A/Title: Purification, characterization and molecular cloning of human hepatic lysosomal
A/Reference number: S41408; MUID:94155897; PMID:8112342

A/Accession: S41408

A/Status: preliminary

A/Molecule type: mRNA

A/Residues: 1-399 <AME>

A/Cross-references: EMBL:X76488; NID:G434305; PIDN:CAA54026.1; PID:G434306

R/Anderson, R.A.; Sando, G.N.

J. Biol. Chem. 266, 22479-22484, 1991

A/Title: Cloning and expression of cDNA encoding human lysosomal acid lipase/cholesteryl

A/Reference number: A39315; MUID:92042192; PMID:1718995

A/Accession: A39315

A/Status: preliminary

A/Molecule type: mRNA

A/Residues: 1-15, 'P', 17-399 <AND>

A/Cross-references: GB:M74775; NID:G187151; PIDN:AAA59519.1; PID:G187152

R/Du, H.; Gregory, G.A.

submitted to the EMBL Data Library, April 1994

A/Description: Structural conservation of putative functional motifs between mouse and h

A/Reference number: S47187

A/Accession: S47187

A/Status: preliminary

A/Molecule type: mRNA

A/Residues: 1-22, 'R', 24-399 <DUH>

A/Cross-references: EMBL:Z31690; NID:G506430; PIDN:CAA83495.1; PID:G506431

C/Genetics:

A/Gene: GDB:LIPA

A/Cross-references: GDB:120153; OMIM:278000

A/Map position: 10q24-10q25

C/Superfamily: triacylglycerol lipase, lingual

C/Keywords: carboxylic ester hydrolase; Glycoprotein

Query Match 57.4%; Score 1290; DB 2; Length 399;

Best Local Similarity 61.0%; Pred. No. 1.4e-105;

Matches 242; Conservative 59; Mismatches 94; Indels 2; Gaps 1;

QY 15 MEMWLLILVAYFQRNVNSVHPTK--AVDPFANVISEIIHQGPCEEEVATEDGYI 72

Db 1 MKMRFLGLVVCILVTLHSEGGGKLTAVDPETNNVSEIIISYWGFPSEBEVATEDGYI 60

QY 73 LSVNRIPRGLVQPKTKGRPVVLLQHLGVGGASNWSINLPNNSLGLFILLADAGFDVWMGNS 132

Db 61 LCLNRPKGRKNSDKGPKPVVFLQHLGLADSSNVTNLANSSGLFILLADAGFDVWMGNS 120

30 105.5 4.7 304 2 A89905 hypothetical prote
31 105 4.7 430 2 T23001 hypothetical prote
32 103.5 4.6 798 2 P96714 probable protease
33 101.5 4.5 929 2 JH0262 Qutr protein - Eme
34 101 4.5 436 2 JCS021 platelet-activatin
35 99.5 4.4 286 2 D75217 probable 2-acetyl-
36 99.5 4.4 1280 2 B34087 hypothetical prote
37 99 4.4 318 2 A82968 ABC transporter, m
38 99 4.4 F98314 unknown permease
39 99 4.4 460 2 G96764 unknown protein F2
40 97.5 4.3 322 2 B70957 probable epha prot
41 97.5 4.3 456 2 E89391 hypothetical prote
42 96.5 4.3 287 1 E64053 hypothetical prote
43 96.5 4.3 1275 2 I38588 reverse transcript
44 96 4.3 809 2 S67153 probable membrane
45 96 4.3 934 2 B82364 DNA polymerase I V

QY 133 RGNWSRKHKTLSDQDEFWAFSDYDEWARFDLPVAVINFILOKTOGQEKIYVYGVSGQTWVG 192
 Db 121 RGNWSRKHKTLSDQDEFWAFSDYDEWARFDLPVAVINFILOKTOGQEKIYVYGVSGQTWVG 180
 QY 193 FIAFTWPELAQIKMYFALAPATATVKAHSPGTEKLLPDMWIKGLFGKKEFLYQTRFL 252
 Db 181 FIAFSQIPELAKRIKMFALGPVAVAFCTSPWAKLGRPLDHLIKDLFGDKEFLPQSAFL 240
 QY 253 RQVIVYLCGVILDOICNSIMLLGGPNTNMMNSRASVYAHTLAGTSVQNTILHWSQAV 312
 Db 241 KWLQTHVCHTHVILKELGNCNLCFLCGFNERNLMSRDVYTHSPAGTSVQNTILHWSQAV 300
 QY 313 NSGELRAFDSGSETKLEKCNQPTPVRYRVRDTPVAMTGGQDMLSNPDEPKMLLSV 372
 Db 301 KFQFQAFDGGSSAKNYFHYNQSPPTNYKMLVETAVWSGGHDLADVDVNWILLTQI 360
 QY 373 TNLVYHKNIPWAHVDFIWLGDAPHMYNEIHLMOQ 409
 Db 361 TNLVYHKNIPWAHVDFIWLGDAPHMYNEIHLMOQ 397

RESULT 2

G01416

lysosomal acid lipase - human

C:Species: Homo sapiens (man)

C>Date: 21-Dec-1996 #sequence_revision 06-Jun-1997 #text_change 18-Jun-1999

C:Accession: G01416

R:Du, H.

submitted to the EMBL Data Library, April 1994

A:Reference number: G06919

A:Accession: G01416

A:Status: preliminary; translated from GB/EMBL/DBJ

A:Molecule type: mRNA

A:Residues: 1-399 <DUX>

A:Cross-references: EMBL:U08464; NID:g505052; PIDN:AAB60328.1; PID:g505053

C:Superfamily: triacylglycerol lipase, lingual

C:Keywords: Glycoprotein

Query Match 57.4%; Score 1290; DB 2; Length 399;
 Best Local Similarity 61.0%; Pred. No. 1.4e-105;
 Matches 242; Conservative 59; Mismatches 94; Indels 2; Gaps 1;

QY 15 MEWMLLILVAYMFQNRVNSVEMPTK--AVDPEAFMNISEIIHQGYPCSEYEYVATEDGYI 72
 Db 1 MKRFLGLVCLVLTWLSHSGSRGLTAVDPETNMNVSEIISYWGPPSEYLVETEDGYI 60
 QY 73 LSVNRI PRGLVQPKKTSRPPVLLQHLGVGASNWSINLNNLSGLTADAGFDVWVWNS 132
 Db 61 LCLNRI PHGRKNHSDKGPVWFLQHLGLADSSNWNVTNLANSLSGLTADAGFDVWVWNS 120
 QY 133 RGNWSRKHKTLSDQDEFWAFSDYDEWARFDLPVAVINFILOKTOGQEKIYVYGVSGQTWVG 192
 Db 121 RGNWSRKHKTLSDQDEFWAFSDYDEWARFDLPVAVINFILOKTOGQEKIYVYGVSGQTWVG 180

QY 193 FIAFTWPELAQIKMYFALAPATATVKAHSPGTEKLLPDMWIKGLFGKKEFLYQTRFL 252
 Db 181 FIAFSQIPELAKRIKMFALGPVAVAFCTSPWAKLGRPLDHLIKDLFGDKEFLPQSAFL 240
 QY 253 RQVIVYLCGVILDOICNSIMLLGGPNTNMMNSRASVYAHTLAGTSVQNTILHWSQAV 312
 Db 241 KWLQTHVCHTHVILKELGNCNLCFLCGFNERNLMSRDVYTHSPAGTSVQNTILHWSQAV 300
 QY 313 NSGELRAFDSGSETKLEKCNQPTPVRYRVRDTPVAMTGGQDMLSNPDEPKMLLSV 372
 Db 301 KFQFQAFDGGSSAKNYFHYNQSPPTNYKMLVETAVWSGGHDLADVDVNWILLTQI 360
 QY 373 TNLVYHKNIPWAHVDFIWLGDAPHMYNEIHLMOQ 409
 Db 361 TNLVYHKNIPWAHVDFIWLGDAPHMYNEIHLMOQ 397

RESULT 3

LIRIT

triacylglycerol lipase (EC 3.1.1.3) precursor, lingual - rat

C:Species: Rattus norvegicus (Norway rat)

C>Date: 30-Sep-1997 #sequence_revision 30-Sep-1987 #text_change 11-Jun-1999

C:Accession: A23045

R:Doeherty, A.J.P.; Bodmer, M.W.; Angal, S.; Verger, R.; Riviere, C.; Lowe, P.A.; Lyons, Nucleic Acids Res. 13, 1891-1903, 1985

A>Title: Molecular cloning and nucleotide sequence of rat lingual lipase cDNA.

A:Reference number: A23045; MUID:85215587; PMID:3839077

A:Accession: A23045

A:Molecule type: mRNA

A:Residues: 1-395 <DOC>

A:Cross-references: GB:X02309; NID:g56595; PIDN:CRA26179.1; PID:g56596

A:Experimental source: strain Sprague-Dawley

A>Note: The partial sequence of the mature protein from a different, unspecified strain

C:Comment: This acid-stable lipase is secreted by the serous (von Ebner's) glands at the

C:Superfamily: triacylglycerol lipase, lingual

C:Keywords: carboxylic ester hydrolase; glycoprotein; lipid digestion; saliva; serous gl

F11-18/Domain: signal sequence #status predicted <SIG>

F19-395/Product: triacylglycerol lipase, lingual #status predicted <MPT>

F19-395/Product: triacylglycerol lipase, lingual #status predicted <MPT>

F19-395/Product: triacylglycerol lipase, lingual #status predicted <MPT>

F19-395/Product: triacylglycerol lipase, lingual #status predicted <MPT>

F19-395/Product: triacylglycerol lipase, lingual #status predicted <MPT>

F19-395/Product: triacylglycerol lipase, lingual #status predicted <MPT>

F19-395/Product: triacylglycerol lipase, lingual #status predicted <MPT>

F19-395/Product: triacylglycerol lipase, lingual #status predicted <MPT>

F19-395/Product: triacylglycerol lipase, lingual #status predicted <MPT>

F19-395/Product: triacylglycerol lipase, lingual #status predicted <MPT>

F19-395/Product: triacylglycerol lipase, lingual #status predicted <MPT>

F19-395/Product: triacylglycerol lipase, lingual #status predicted <MPT>

F19-395/Product: triacylglycerol lipase, lingual #status predicted <MPT>

F19-395/Product: triacylglycerol lipase, lingual #status predicted <MPT>

F19-395/Product: triacylglycerol lipase, lingual #status predicted <MPT>

F19-395/Product: triacylglycerol lipase, lingual #status predicted <MPT>

F19-395/Product: triacylglycerol lipase, lingual #status predicted <MPT>

F19-395/Product: triacylglycerol lipase, lingual #status predicted <MPT>

F19-395/Product: triacylglycerol lipase, lingual #status predicted <MPT>

F19-395/Product: triacylglycerol lipase, lingual #status predicted <MPT>

F19-395/Product: triacylglycerol lipase, lingual #status predicted <MPT>

F19-395/Product: triacylglycerol lipase, lingual #status predicted <MPT>

F19-395/Product: triacylglycerol lipase, lingual #status predicted <MPT>

F19-395/Product: triacylglycerol lipase, lingual #status predicted <MPT>

F19-395/Product: triacylglycerol lipase, lingual #status predicted <MPT>

F19-395/Product: triacylglycerol lipase, lingual #status predicted <MPT>

F19-395/Product: triacylglycerol lipase, lingual #status predicted <MPT>

F19-395/Product: triacylglycerol lipase, lingual #status predicted <MPT>

F19-395/Product: triacylglycerol lipase, lingual #status predicted <MPT>

F19-395/Product: triacylglycerol lipase, lingual #status predicted <MPT>

F19-395/Product: triacylglycerol lipase, lingual #status predicted <MPT>

F19-395/Product: triacylglycerol lipase, lingual #status predicted <MPT>

F19-395/Product: triacylglycerol lipase, lingual #status predicted <MPT>

F19-395/Product: triacylglycerol lipase, lingual #status predicted <MPT>

F19-395/Product: triacylglycerol lipase, lingual #status predicted <MPT>

F19-395/Product: triacylglycerol lipase, lingual #status predicted <MPT>

F19-395/Product: triacylglycerol lipase, lingual #status predicted <MPT>

F19-395/Product: triacylglycerol lipase, lingual #status predicted <MPT>

F19-395/Product: triacylglycerol lipase, lingual #status predicted <MPT>

F19-395/Product: triacylglycerol lipase, lingual #status predicted <MPT>

F19-395/Product: triacylglycerol lipase, lingual #status predicted <MPT>

F19-395/Product: triacylglycerol lipase, lingual #status predicted <MPT>

F19-395/Product: triacylglycerol lipase, lingual #status predicted <MPT>

F19-395/Product: triacylglycerol lipase, lingual #status predicted <MPT>

F19-395/Product: triacylglycerol lipase, lingual #status predicted <MPT>

F19-395/Product: triacylglycerol lipase, lingual #status predicted <MPT>

F19-395/Product: triacylglycerol lipase, lingual #status predicted <MPT>

F19-395/Product: triacylglycerol lipase, lingual #status predicted <MPT>

F19-395/Product: triacylglycerol lipase, lingual #status predicted <MPT>

F19-395/Product: triacylglycerol lipase, lingual #status predicted <MPT>

F19-395/Product: triacylglycerol lipase, lingual #status predicted <MPT>

F19-395/Product: triacylglycerol lipase, lingual #status predicted <MPT>

F19-395/Product: triacylglycerol lipase, lingual #status predicted <MPT>

F19-395/Product: triacylglycerol lipase, lingual #status predicted <MPT>

F19-395/Product: triacylglycerol lipase, lingual #status predicted <MPT>

F19-395/Product: triacylglycerol lipase, lingual #status predicted <MPT>

F19-395/Product: triacylglycerol lipase, lingual #status predicted <MPT>

F19-395/Product: triacylglycerol lipase, lingual #status predicted <MPT>

F19-395/Product: triacylglycerol lipase, lingual #status predicted <MPT>

F19-395/Product: triacylglycerol lipase, lingual #status predicted <MPT>

F19-395/Product: triacylglycerol lipase, lingual #status predicted <MPT>

F19-395/Product: triacylglycerol lipase, lingual #status predicted <MPT>

F19-395/Product: triacylglycerol lipase, lingual #status predicted <MPT>

F19-395/Product: triacylglycerol lipase, lingual #status predicted <MPT>

Db 111 GNRGNTYNNKXKOLXPSHSAFWDSDENATYDNLAMINHVLEVTGQDSVYMGHSQOT 170
 QY 190 TMGFIAPSTWP-ELAQIKIKYFALAPATVYKAKSGPTKFLLLPDMMIKG---LFGKKEP 245
 Db 171 LTMFSLKSDGGSFAKKIKKFFALAPIGSVKIKGFLSPFANYFSLFDFGWFIDFAGGEF 230
 QY 246 LYQTRFLRQLVYLCCGVILD-QICSNIMLLGGFNNTNMNSRASVVAHAHTLAGTSVON 304
 Db 231 LNNWAKMAKADICGLKVEADLNDNVLFLAGPESDQWQTRVPVATHPAGTSTQN 290
 QY 305 ILHWSQAVNSGELRAFDSWSETKNLEKCNQPTFVRYRVRDMT---VPTAMWTGGQDMLS 360
 Db 291 IVEHMQVMVHHGGVPAYDWGKT-KKKYGOANPPEY---DFTAKGTKIYLYWSADMLA 346
 QY 361 NPEDV-KMLLSEVNTLIYHKN--IPWAHVDFITGLDAPHRMNEIHL 406
 Db 347 DTPDVPDYLLTRLNPALVAQNHLDPDYNHLDFTWGLRAPDDIYRPAKL 395

RESULT 7
 T22290
 hypothetical protein F46B6.8 - Caenorhabditis elegans
 C:Species: Caenorhabditis elegans
 C:Date: 15-Oct-1999 #sequence_revision 15-Oct-1999 #text_change 21-Jan-2000
 C:Accession: T22290
 R:McMurray, A.
 submitted to the EMBL Data Library, April 1996
 A:Reference number: Z19542
 A:Accession: T22290
 A:Status: preliminary; translated from GB/EMBL/DBJ
 A:Molecule type: DNA
 A:Residues: 1-411 <MIL>
 A:Cross-references: EMBL:Z70780; PIDN:CAA94824.1; GSPDB:GN00023; CESP:F46B6.8
 A:Experimental source: clone F46B6
 C:Genetics:
 A:Gene: CESP:F46B6.8
 A:Map position: 5
 A:Introns: 35/3; 69/1; 104/1; 226/3
 C:Superfamily: triacylglycerol lipase, lingual

Query Match 33.4%; Score 751; DB 2; Length 411;
 Best Local Similarity 40.4%; Pred. No. 4.2e-58;
 Matches 163; Conservative 72; Mismatches 138; Indels 30; Gaps 12;
 QY 20 LILVAYMFQNVNSVHMTKAVDPEAFNISEIIHQGYPCPEYEVEATEDGYILS 79
 Db 10 LICTSFLLQITINA-HPDD---DPELNMNTSQIIRWGYKAEVHTVTEDGYILOMCRIP 65
 QY 80 RGLVQPKKT-----GSRPVLLQHLVGGASNINLNNLSGLFILDAGFDVMMGNSR 133
 Db 66 YG-----KTSVTPNGKRPVILLQGLLACASDWNILPTQSAFVPADAGFDVWLGVR 120
 QY 134 GNWSRKHKTLSDQDEFWAFSDYEMARFDLPVAVINFILOKTQEKIYVYVYSGTGMF 193
 Db 121 GTTYGRKNTKLDPSSETAFWQFSWDEMAQYDVPAMVDHVLAMTQENLYTMGHSQGTLMF 180
 QY 194 --TAFSTWPELAKIKMYFALAPATVYKAKSGPTKFLLLPDMMIKG---LFGKKEFLYQ 248
 Db 181 THLAKDTGSGFAKKIKKFFALAPIGAVKNTKFLPSYFAHKSPFEPDGTWELFSGKDFLPD 240
 QY 249 TRFLRQLVYLCCGVILD-QICSNIMLLGGFNNTNMNSRASVVAHAHTLAGTSVONILH 307
 Db 241 NWITMAAKDICGASEAEALCNELFLIGPESDQWNASRTAIYSSQDPAGTSTQNIHV 300
 QY 308 WSAVNSGELRAFDSWSETKNLEKCNQPTFVRY---RVDMTVPYATMTGQDMLSNP 364
 Db 301 WMQVNRGRVPAPDWGKKI-KKKYQGDTPPEYDFGAIKGTKI--HLYNSDDDDMLGDPDT 357
 QY 365 V-KMLLSEVNTLIYHKNI--PEWAHVDFITGLDAPHRMNEI 404
 Db 358 IHDPELLKXELAPAVIAENVNLEKYNHLDTWGLNATFQIYDKAI 400

RESULT 8

T22675
 hypothetical protein F54F3.3 - Caenorhabditis elegans
 C:Species: Caenorhabditis elegans
 C:Date: 15-Oct-1999 #sequence_revision 15-Oct-1999 #text_change 21-Jan-2000
 C:Accession: T22675
 R:Percy, C.; Lloyd, C.
 submitted to the EMBL Data Library, September 1996
 A:Reference number: Z19598
 A:Accession: T22675
 A:Status: preliminary; translated from GB/EMBL/DBJ
 A:Molecule type: DNA
 A:Residues: 1-405 <MIL>
 A:Cross-references: EMBL:Z79696; PIDN:CAB01973.1; GSPDB:GN00023; CESP:F54F3.3
 A:Experimental source: clone F54F3
 C:Genetics:
 A:Gene: CESP:F54F3.3
 A:Map position: 5
 A:Introns: 31/3; 65/1; 274/3; 366/3
 C:Superfamily: triacylglycerol lipase, lingual

Query Match 32.5%; Score 730; DB 2; Length 405;
 Best Local Similarity 38.6%; Pred. No. 2.9e-56;
 Matches 161; Conservative 69; Mismatches 155; Indels 32; Gaps 11;
 QY 15 MEMMLLILVAYMFQNVNSVHMTKAVDPEAFNISEIIHQGYPCPEYEVEATEDGYILS 74
 Db 1 MRSWSTVMLAVLATAATVFGH---DADPEMKMTTPQIIMRWGYPAMIYDVTTEDGYILE 56
 QY 75 VNRIPRGLVQPKKT-----GSRPVLLQHLVGGASNINLNNLSGLFILDAGFDVW 128
 Db 57 LHRIPYG-----KTNVTPNGKRPVFMQGLSCSSNNVVNLTESAFLPADAGYDWM 111
 QY 129 MGNRGNNAWRKHKTLSDQDEFWAFSDYEMARFDLPVAVINFILOKTQEKIYVYVYSG 188
 Db 112 LGNFRGNTYSKCHKNLKXPSHSAFWDSDENATYDNLAMIEKALEVTGQDSLYYIGHSG 171
 QY 189 T-TMGFIAPSTWP-----LAQIKMYFALAPATVYKAKSGPTKFLLLPDMMIKG---LF 240
 Db 172 TLTM-----FSRLSEDKVGNKIKKFFALAPVGSVKHKGKALFFADYSLFDFGWFDFV 227
 QY 241 GKKEFLYQTRFLRQLVYLCCGVILDQICSNIMLLGGFNNTNMNSRASVVAHAHTLAG 299
 Db 228 GSGEFLNNIMKLVSVCAGLKVRAVCDDVWELIAGPESQNLNATRPVIVVAHTPAG 287
 QY 300 TSVONILHWSQAVNSGELRAFDSWSETKNLEKCNQPTFVRYRVRDMTVPYATMTGQDML 359
 Db 288 TSTQNIYHWTQMVNRHGGTPKYDYG-EKGNKKHYGOANVPAYDFTTVNRPVYLYWGSDDL 346
 QY 360 SNPEDV-KMLLSEV--TNLIYHKNIPWAHVDFITGLDAPHRMNEIHLMOQEBTN 413
 Db 347 ADPTDVTDFLLTHLNPSTVQNNKLDYNHLDFIWGLRAPKDIYEPIDIVRNDVLN 403

RESULT 9

H88930
 protein R11G11.14 [imported] - Caenorhabditis elegans
 C:Species: Caenorhabditis elegans
 C:Date: 10-May-2001 #sequence_revision 10-May-2001 #text_change 24-May-2001
 C:Accession: H88930
 R:Anonymous, The C. elegans Sequencing Consortium.
 Science 282, 2012-2018, 1998
 A:Title: Genome sequence of the nematode C. elegans: a platform for investigating biology
 A:Reference number: A75000; MUID:99069613; PMID:9851916
 A:Note: see websites genome.wustl.edu/gsc/C.elegans/ and www.sanger.ac.uk/projects/C.ele
 A:Note: published errata appeared in Science 283, 35, 1999; Science 283, 2103, 1999; and
 A:Accession: H88930
 A:Status: preliminary
 A:Molecule type: DNA
 A:Residues: 1-405 <STO>
 A:Cross-references: GB:chr V; PIDN:AAC69088.1; PID:G2384863; GSPDB:GN00023; CESP:R11G11.14
 A:Note: Similar to lipase, R11G11.14

RESULT 12

JT0949
 egg-specific protein - silkworm
 C:Species: Bombyx mori (silkworm)
 C:Date: 31-Mar-1992 #sequence_revision 31-Mar-1992 #text_change 26-Feb-1998
 C:Accession: JT0949; A28527
 R:Sato, Y.; Yamashita, O.
 submitted to JIPID, September 1991
 A:Reference number: JT0949
 A:Accession: JT0949
 A:Molecule type: DNA
 A:Residues: 1-559 <SAT>
 A:Experimental source: larva
 A>Note: this protein is a homotrimer
 R:Indrasith, L.S.; Sasaki, T.; Yamashita, O.
 J. Biol. Chem. 263, 1045-1051, 1988
 A:Title: A unique protease responsible for selective degradation of a yolk protein in B.
 A:Reference number: A28527; MUID:88087166; PMID:3275655
 A:Accession: A28527
 A:Molecule type: protein
 A:Residues: 13-40/133-144, 'T', 146-152/229-248 <IND>
 C:Genetics:
 A:Map position: 19
 C:Keywords: egg yolk; homotrimer
 F:132-133/Cleavage site: Lys-Asn (egg 30.5K cysteine proteinase) #status experimental
 F:228-229/Cleavage site: Arg-Asp (egg 30.5K cysteine proteinase) #status experimental

Query Match 25.6%; Score 574.5; DB 2; Length 559;
 Best Local Similarity 33.2%; Pred. No. 2.3e-42;
 Matches 132; Conservative 82; Mismatches 152; Indels 31; Gaps 9;

QY 32 NSVH---MPTKAVDPEAF-WNISEIIOHQGPCEEEVATEDGYIISVNRIPRGLVQPKK 87
 DB 171 NTFDAISETQRENEDFHLNATELLKKHQYVEEHTVATDDGYHITVLRIIP-PTHQTRD 229
 QY 88 TGSFVVLLOHGLVGGASNWSLNPNNSLGFILADAGFDVWNGNSRGNASRKHKTLSID 147
 DB 230 DKKFPVALLHGLLSADDLLMGPSKSLAYMLSDAGYDVLGNVGNKYSRSHVSKHPA 289
 QY 148 QDEFWAFSDYEMARFDLPVAVINFLQTKGQEKIYYVYSGQTTMGTFIAFTMPPELAQKIK 207
 DB 290 LNDFWKFSNDEIALHDLPAIDHVDLSGQERLHYHIGHSQCATTFFALMSEQPSYNEKIV 349
 QY 208 MYFALAPIATYKHAKS-----PGRKFL-LLPDMTKLGFCKKFTLYQTRFLRLQVLYL 259
 DB 350 SMHALSPVYNNYVRSPLFRMIAPTSKPYQVIHDDQVGHGAFEPGKHLIET-----F 400
 QY 260 CGQVILDQ-----ICSNIMLLGGFTNNNNMNRASVYAAHTLAGTSVQVNIHWSQAVNS 314
 DB 401 GGAACREKLGCRHYCNLNLYVIGINVYNQDADIVPVVMAHLPAGTSARVMKQYGNVAS 460
 QY 315 GELRAPDWSGTXLEKCNQPTPVYRVDRMTVPTAMWTGGQDWLSNPEDVYKMLSEVTN 374
 DB 461 HDPKRYNGAET-NMKYVGASEPPSYDLKSVAPVNLVYHSHDADLAPKDKVEXLQENLPN 519
 QY 375 LIYHKNIPWAH---VDFIWGLDAPHRYNIEIHLMQ 408
 DB 520 VKQSFVEPEQHFDDLFQFSKAPDIVYQKLMENMQ 556

RESULT 13
 T39540
 triglyceride lipase-cholesterol esterase - fission yeast (Schizosaccharomyces pombe)
 C:Species: Schizosaccharomyces pombe
 C:Date: 03-Dec-1999 #sequence_revision 03-Dec-1999 #text_change 17-Mar-2000
 C:Accession: T39540
 R:Wood, V.; Rajandream, M.A.; Barrell, B.G.; Beck, A.; Reinhardt, R.
 submitted to the EMBL Data Library, February 1998
 A:Reference number: Z21862
 A:Accession: T39540
 A>Status: preliminary; translated from GB/EMBL/DBJ
 A:Molecule type: DNA

A:Residues: 1-443 <WOO>
 A:Cross-references: EMBL:AL021748; PIDN:CAAL6863.1; GSPDB:GNO0067; SPDB:SPBC16A3.12c
 A:Experimental source: strain 972h-; cosmid c16a3
 C:Genetics:
 A:Gene: SPDB:SPBC16A3.12c
 A:Map position: 2
 A:Introns: 11/2; 152/3; 282/1; 398/3
 C:Superfamily: triacylglycerol lipase, lingual

Query Match 22.5%; Score 505; DB 2; Length 443;
 Best Local Similarity 29.6%; Pred. No. 2.1e-36;
 Matches 125; Conservative 91; Mismatches 154; Indels 52; Gaps 13;

QY 9 WIVSHRMEMW--LLILVAYMFORNVNSVHMPTKAVDPEAFMNI-SBIIHQGPCEYEVA 66
 DB 35 YMIPEVRVWFVRVRIIVLYHYIYSSKTTDGTGMDAV--QKCRNIYEICEAFGYRVEEHLVR 92
 QY 67 TEDGYIISVNRIPRGLVQPKKT-GSRPVLLQHLVGGASNWSL-NLPNNSLGFILADAG 124
 DB 93 TDNFIILCHRI-----THPKSQSHKREVYVYCHHGLMTNSLWVAVNESRSLFPVLIESG 148
 QY 125 FDVWNGNSRGNASRKHKTLSIDQDEFWAFSDYEMARFDLPVAVINFLQTKGQEKIYYVG 184
 DB 149 YDWLGNNGNKNYSRKHITVYKPKDEEFWFSLDMDAMFDIPDTVDYILRETGREKLNYYIG 208
 QY 185 YSQGTTMGTFIAFTMPPELAQKIKMYFALAPIATYKHAKSPTGKFL--PDMMLKGLFGK 242
 DB 209 FSQGTACAMALSNPDLNDKNWIFIGLAPAYAPKGSFYVDYIVKVNPKIMYH-LFGR 267
 QY 243 KEFLYQTRFLRLQVLYLGGVILDOICSNIMLLGGFTNNNNMNRASVYAAHTLAGTSV 302
 DB 268 RCLIPSVTFWQV----ICYPPIFVKIVDVSILKLFNWDLSNLSINQKLCGYAHLYSFSSV 323
 QY 303 QNHLWSQAVNSGELRAPD-----WGSE-----TNLEKCNQPTPVYRVDRMT 346
 DB 324 KSVVFWLQIIKXCTFQLYDMDALLAGYGRHYQVLPFTNNI-KC-----368
 QY 347 VPTAMWTGGQDWLSNPEDVYKMLSEVTNLYHKNIPWAHVDFTWGLDAPHRYMNEIHL 406
 DB 369 -PMILWGGKDTLLNMEVMTALPPHAKEV---SIAHYEHLDFLWGDVKEEVPVVIDA 424
 QY 407 MQ 408
 DB 425 LK 426

RESULT 14

T31611
 hypothetical protein Y508A.g - Caenorhabditis elegans
 C:Species: Caenorhabditis elegans
 C:Date: 29-Oct-1999 #sequence_revision 29-Oct-1999 #text_change 29-Oct-1999
 C:Accession: T31611
 R:Steward, C.
 submitted to the EMBL Data Library, September 1999
 A:Reference number: Z21047
 A:Accession: T31611
 A>Status: preliminary; translated from GB/EMBL/DBJ
 A:Molecule type: DNA
 A:Residues: 1-1585 <WIL>
 A:Cross-references: EMBL:AL117200; NID:e1549770; PIDN:CAB55050.1; CBSP:Y508A.g
 A:Experimental source: clone Y508A
 C:Genetics:
 A:Gene: CBSP:Y508A.g
 A:Introns: 25/3; 60/1; 133/2; 217/3; 270/3; 337/2; 400/1; 746/2

Query Match 21.3%; Score 478.5; DB 2; Length 1585;
 Best Local Similarity 30.6%; Pred. No. 2.8e-33;
 Matches 118; Conservative 67; Mismatches 172; Indels 29; Gaps 9;

QY 44 EAFNNISEIIOHQGPCEYEVEATEDGYIISVNRIPRGLVQPKKTG--SRPVLLQHLGLV 101
 DB 20 EETLDAADATISHGYVTEKHVYTTDDGYTVQLQRIIPVGRDRSILGSKRPVVFMEGLF 79

GenCore version 5.1.6
Copyright (c) 1993 - 2004 Compugen Ltd.

OM protein - protein search, using sw model

Run on: February 19, 2004, 14:16:19 ; Search time 9.88561 Seconds
(without alignments)
2012.248 Million cell updates/sec

Title: US-10-042-431-47

Perfect score: 2247

Sequence: 1 MLETLRSQWIVSHRMWELL.....IHLMQEETNLSQRCCEAVL 423

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 127863 seqs, 47026705 residues

Total number of hits satisfying chosen parameters: 127863

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : SwissProt_41.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	1289	57.4	399	1 LICH_HUMAN	P38571 homo sapien
2	1220.5	54.3	397	1 LICH_RAT	Q84194 rattus norv
3	1197.5	53.3	397	1 LICH_MOUSE	Q520M5 mus musculu
4	1164.5	51.8	395	1 LIPG_RAT	P80634 rattus norv
5	1154	51.4	398	1 LIPG_CANFA	P80035 canis famil
6	1139	50.7	397	1 LIPG_BOVIN	Q29458 bos taurus
7	1119	49.8	398	1 LIPG_HUMAN	P07098 homo sapien
8	663	29.5	394	1 LIP3_DROME	O46108 drosophila
9	546	24.3	439	1 LIP1_DROME	O46107 drosophila
10	422.5	18.8	548	1 TGL1_YEAST	P34163 saccharomyc
11	126	5.6	987	1 YD94_METUA	Q58789 methanococ
12	105	4.7	430	1 ORC2_CABEL	Q21037 caenorhabdi
13	101	4.5	436	1 PAPA_CAVPO	P70683 c platelet-
14	96.5	4.3	287	1 Y193_HAETN	Q57427 haemophilus
15	95.5	4.3	1259	1 LIN1_HUMAN	P08547 homo sapien
16	93.5	4.2	278	1 PRXC_STRAU	O31168 streptomyce
17	93	4.1	569	1 PRD_FLAPA	Q08210 plasmodium
18	92	4.1	748	1 PTIP_ECOLI	P37177 escherichia
19	92	4.1	748	1 PTIP_SALTY	P37178 salmonella
20	91	4.0	370	1 P2C2_SCHPO	Q09172 schizosacch
21	90.5	4.0	1037	1 ACRD_ECOLI	P24177 escherichia
22	90	4.0	277	1 BPA2_STRAU	P29715 streptomyce
23	90	4.0	1122	1 Y3C3_YEAST	P33280 saccharomyc
24	90	4.0	1136	1 POLG_OMV	Q9esd7 mus musculu
25	90	4.0	2083	1 DYSP_MOUSE	P02023 o genome po
26	89	4.0	330	1 EXOA_RHINE	P33691 rhizobium m
27	89	4.0	422	1 YG83_SHEON	P59352 shewanella
28	89	4.0	521	1 ACH1_SCHPO	Q9uni9 schizosacch
29	88.5	3.9	185	1 YZTH_BACSU	O31432 bacillus su
30	88.5	3.9	869	1 CFAC_ECOLI	P25733 escherichia
31	87	3.9	473	1 YDRH_LACIA	Q04937 lactococcus
32	86.5	3.8	382	1 YDRH_HAETN	P44861 haemophilus
33	86.5	3.8	530	1 SYN_ANASP	Q8z068 anabaena sp

RESULT 1

ID	LICH_HUMAN	STANDARD;	PRT;	399 AA.
AC	P38571: Q16529; Q96EJ0;			
DT	01-OCT-1994 (Rel. 30, Created)			
DT	01-OCT-1994 (Rel. 30, Last sequence update)			
DT	15-SEP-2003 (Rel. 42, Last annotation update)			
DE	Lysosomal acid lipase/cholesterol ester hydrolase precursor			
DE	(EC 3.1.1.13) (LAL) (Acid cholesteryl ester hydrolase) (Sterol			
DE	esterase) (Lipase A) (Cholesteryl esterase).			
GN	LIPA.			
OS	Homo sapiens (Human).			
OC	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;			
OC	Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.			
OX	NCBI_TaxID=9606;			
[1]	SEQUENCE FROM N.A., AND SEQUENCE OF 196-212; 277-297 AND 305-315.			
RN	MEDLINE=92042192; PubMed=1718995;			
RX	Anderson R.A., Sando G.N.;			
RA	"Cloning and expression of cDNA encoding human lysosomal acid			
RT	lipase/cholesterol ester hydrolase. Similarities to gastric and			
RT	lingual lipases."			
RL	J. Biol. Chem. 266:22479-22484(1991).			
[2]	SEQUENCE FROM N.A.			
RN	MEDLINE=94155897; PubMed=8112342;			
RP	Ameis D., Merkel M., Eckerskorn C., Greten H.;			
RX	"Purification, characterization and molecular cloning of human			
RA	hepatic lysosomal acid lipase."			
RT	Eur. J. Biochem. 219:905-914(1994).			
[3]	SEQUENCE FROM N.A.			
RN	MEDLINE=96363957; PubMed=8725147;			
RP	Du H., Witte D.P., Grabowski G.A.;			
RX	"Tissue and cellular specific expression of murine lysosomal acid			
RA	lipase mRNA and protein."			
RT	J. Lipid Res. 37:937-949(1996).			
[4]	SEQUENCE FROM N.A.			
RN	MEDLINE=22388257; PubMed=12477932;			
RP	Strausberg R.L., Feingold E.A., Grouse L.H., Derge J.G.,			
RX	Klaunig R.D., Collins F.S., Wagner L., Shenmen C.M., Schuler G.D.,			
RA	Altschul S.F., Zeeberg B., Buetow K.H., Schaefer C.F., Shat N.K.,			
RT	Hopkins R.F., Jordan H., Moore T., Max S.I., Wang J., Hsieh F.,			
RL	Diatchenko L., Marsina K., Farmer A.A., Rubin G.M., Hong L.,			
RA	Stapleton M.J., Soares M.B., Donald M.F., Casavant T.L., Schetz T.B.,			
RX	Brownstein M.J., Usdin T.B., Toshiyuki S., Carninci P., Prange C.,			
RA	Raha S.A., Loquellano N.A., Peters G.J., Abramson R.D., Mullaly S.J.,			
RX	Bosak S.A., McEwan P.J., McKernan K.J., Malek J.A., Gunaratne P.H.,			
RA	Richards S., Worley K.C., Hale S., Garcia A.M., Gay L.J., Rulyk S.W.,			
RX	Villalon D.K., Muzny D.C., Hale S., Sodergren E.J., Lu X., Gibbs R.A.,			
RA	Fahey J., Helton E., Kettelman M., Madan A., Rodrigues S., Sanchez A.,			
RL	Whitting M., Madan A., Young A.C., Shevchenko Y., Bouffard G.G.,			

ALIGNMENTS

34	86.5	3.8	685	1	INVA_SALTY
35	86.5	3.8	1882	1	Y468_MYCPN
36	86	3.8	525	1	ACUB_NEUCR
37	86	3.8	1866	1	VGNB_CPMV
38	85.5	3.8	428	1	YB63_SCHPO
39	85	3.8	292	1	FAEB_NEUCR
40	85	3.8	429	1	ARB2_ECOLI
41	85	3.8	429	1	EIOA_BUCBP
42	85	3.8	837	1	UPBA_DICDI
43	85	3.8	4128	1	PRKD_MOUSE
44	84.5	3.8	875	1	AMD2_XENLA
45	84	3.7	451	1	YB27_YEAST

P35657	salmonella
P5109	mycoplasma
P51937	neurospora
P03600	cowpea mosa
O14249	schizosacch
Q9hgr3	neurospora
P52146	escherichia
Q89ak4	buchnera ap
Q89ak4	buchnera ap
P54201	dictyosteli
P97313	mus musculu
P12890	xenopus lae
P38295	saccharomyc

RA Blakesley R.W., Touchman J.W., Green E.D., Dickson M.C.,
 RA Rodriguez A.C., Grimwood J., Schmutz J., Myers R.M.,
 RA Butterfield Y.S.N., Krzywinski M.I., Skalska U., Smalhus D.E.,
 RA Scherch A., Schein J.E., Jones S.J.M., Marra M.A.;
 RT "Generation and initial analysis of more than 15,000 full-length
 human and mouse cDNA sequences.";
 RL Proc. Natl. Acad. Sci. U.S.A. 99:16899-16903(2002).
 RN [5]
 RP VARIANT CDS/WD PRO-200.
 RX MEDLINE=94195814; PubMed=8146180;
 RA Anderson R.A., Byrum R.S., Coates P.M., Sando G.N.;
 RT "Mutations at the lysosomal acid cholesteryl ester hydrolase gene
 locus in Wolman disease.";
 RL Proc. Natl. Acad. Sci. U.S.A. 91:2718-2722(1994).
 RN [6]
 RP VARIANTS CDS ARG-129 AND PRO-129.
 RX MEDLINE=98295576; PubMed=9633819;
 RA Ries S., Buechler C., Schindler G., Aelandis C., Ameis D., Gasche C.,
 RA Jung N., Schambach A., Fehrer P., Vanier M.T., Belli D.C.,
 RA Greten H., Schmitz G.;
 RT "Different missense mutations in histidine-108 of lysosomal acid
 lipase cause cholesteryl ester storage disease in unrelated compound
 heterozygous and hemizygous individuals.";
 RL Hum. Mutat. 12:44-51(1998).
 CC -1- FUNCTION: CRUCIAL FOR THE INTRACELLULAR HYDROLYSIS OF CHOLESTERYL
 ESTERS AND TRIGLYCERIDES THAT HAVE BEEN INTERNALIZED VIA RECEPTOR-
 MEDIATED ENDOCYTOSIS OF LIPOPROTEIN PARTICLES. IMPORTANT IN
 CC MEDIATING THE EFFECT OF LDL (LOW DENSITY LIPOPROTEIN) UPTAKE ON
 CC SUPPRESSION OF HYDROXYMETHYLGLUTARYL-COA REDUCTASE AND ACTIVATION
 CC OF ENDOGENOUS CELLULAR CHOLESTERYL ESTER FORMATION.
 CC -1- CATALYTIC ACTIVITY: A sterol ester + H(2)O = a sterol + a fatty
 acid.
 CC -1- SUBCELLULAR LOCATION: Lysosomal.
 CC -1- DISEASE: DEFECTS IN LIPA ARE THE CAUSE OF THE SEVERE INFANTILE-
 CC ONSET WOLMAN DISEASE (WD) AND THE Milder LATE-ONSET CHOLESTERYL
 CC ESTER STORAGE DISEASE (CESD).
 CC -1- SIMILARITY: BELONGS TO THE ABHYDROLASE SUPERFAMILY. LIPASE FAMILY.
 CC
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@isb-sib.ch).
 CC
 CC -----
 DR EMBL; M74775; AAB59519.1; -;
 DR EMBL; U04285; AAB60327.1; -;
 DR EMBL; U04286; AAB60327.1; JOINED.
 DR EMBL; U04287; AAB60327.1; JOINED.
 DR EMBL; U04288; AAB60327.1; JOINED.
 DR EMBL; U04290; AAB60327.1; JOINED.
 DR EMBL; U04291; AAB60327.1; JOINED.
 DR EMBL; U04292; AAB60327.1; JOINED.
 DR EMBL; U04293; AAB60327.1; JOINED.
 DR EMBL; X76488; CAAS4026.1; -;
 DR EMBL; Z33690; CAAS4026.1; -;
 DR EMBL; U04654; AAB60328.1; -;
 DR EMBL; BC012287; AAH12287.1; -;
 DR PIR; S41408; S41408.
 DR Genew; HGNC:6617; LIPA.
 DR MIM; 278000; -;
 DR GO; GO:0005764; C:lysosome; TAS.
 DR GO; GO:0006487; P:N-linked glycosylation; TAS.
 DR InterPro; IPR000073; A/B hydrolase.
 DR InterPro; IPR000734; Lipase.
 DR InterPro; IPR000379; Ser esters site.
 DR Pfam; PF04083; abhydrolase.1.
 DR Pfam; PF00561; abhydrolase.1.
 DR PROSITE; PS00120; LIPASE SER; 1.
 DR KW Hydrolase; Lipid degradation; Glycoprotein; Signal; Lysosome;
 KW Disease mutation; Polymorphism.
 FT SIGNAL 1 21 POTENTIAL.

FT CHAIN 22 399 LYSOSOMAL ACID LIPASE/CHOLESTERYL ESTER
 FT ACT_SITE 174 174 CHARGE RELAY SYSTEM (BY SIMILARITY).
 FT ACT_SITE 374 374 CHARGE RELAY SYSTEM (BY SIMILARITY).
 FT CARBOHYD 36 36 N-LINKED (GLCNAC. .) (POTENTIAL).
 FT CARBOHYD 72 72 N-LINKED (GLCNAC. .) (POTENTIAL).
 FT CARBOHYD 101 101 N-LINKED (GLCNAC. .) (POTENTIAL).
 FT CARBOHYD 161 161 N-LINKED (GLCNAC. .) (POTENTIAL).
 FT CARBOHYD 273 273 N-LINKED (GLCNAC. .) (POTENTIAL).
 FT CARBOHYD 321 321 N-LINKED (GLCNAC. .) (POTENTIAL).
 FT VARIANT 16 16 P -> T.
 FT VARIANT 129 129 H -> P (IN CESD).
 FT VARIANT 129 129 H -> R (IN CESD).
 FT VARIANT 200 200 L -> P (IN CESD AND WD).
 FT CONFLICT 23 23 G -> R (IN REF. 3).
 FT CONFLICT 29 29 V -> L (IN REF. 4).
 SQ SEQUENCE 399 AA; 45415 MW; 55F27391306B609A CRC64;
 Query Match 57.4%; Score 1289; DB 1; Length 399;
 Best Local Similarity 63.5%; Pred. No. 3.5e-103; Indels 0; Gaps 0;
 Matches 235; Conservative 55; Mismatches 80;
 QY 40 AVDPAPMNISEIIHQGYPCBEYEVATEDGYILSVNRIPIRGVLQPKTKGSRPVLLQGG 99
 DB 28 AVDPETNMVSEIISYWGPPSEYLVETEDGYILNRIPIHGRKNHSDKGPKEVFLQGG 87
 QY 100 LVGASWINLNNISLGFILADAGFDVWNGSRGNANSHKHTLSIDODEWAFSYDEM 159
 DB 88 LLADSSNVTNLANSLSLGFILADAGFDVWNGSRGNANSHKHTLSVQDEFWAFSYDEM 147
 QY 160 ARPLPAVINFILOKQGEKIIYVYVGSQGTMTGFIAPSTMPPELAQIKKYPALAPATVK 219
 DB 148 AKYDLASINFILOKQGEKIIYVYVGSQGTMTGFIAPSTMPPELAQIKKYPALAPATVK 207
 QY 220 HAKSPGTFKLLPDMITKGLFGKKEFYIYQTRQLVYLCGVILDDQICSNMILLGGF 279
 DB 208 FCTSPMAKRLPDLHKLDFGDKFELPQSAFLKWLGTHTVCTHVLKELCGNLCFLCGF 267
 QY 280 NTNNMNSRASVYAAHTAGTSVQNTLHWSOAVNSGELRAFDWGSSEKLEKCNQPTPVR 339
 DB 268 NERNLNRSVDVYTHSPAGTSVQNTLHWSOAVNSGELRAFDWGSSEKLEKCNQPTPVR 327
 QY 340 YRVSDMTVPTAMTGGQDWLSPNEDVKLLSEVNTLIYHKNIPFWAHVDFWGLDAPHRM 399
 DB 328 YNVKMLVPTAVNSGCHDLADVVDVNNILLTQITNLVFEHSIPEWEHLDFWGLDAPWRL 387
 QY 400 YNEIHLMOQ 409
 DB 388 YNKIINLMRK 397
 RESULT 2
 ID LICH RAT STANDARD; PRT; 397 AA.
 AC Q64194;
 DT 28-FEB-2003 (Rel. 41, Created)
 DT 28-FEB-2003 (Rel. 41, Last sequence update)
 DT 28-FEB-2003 (Rel. 41, Last annotation update)
 DE Lysosomal acid lipase/cholesteryl ester hydrolase precursor
 DE (EC 3.1.1.13) (LAL) (Acid cholesteryl ester hydrolase) (Sterol
 DE esterase) (Lipase A) (Cholesteryl esterase).
 GN LIPA OR LAL.
 OS Rattus norvegicus (Rat).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Rattus.
 OX NCBI_TaxID=10116;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC TISSUE=Liver;

RX MEDLINE=96129534; PubMed=8576647;
 RA Nakagawa H., Matsubara S., Kuriyama M., Yoshidome H., Fujiyama J.,
 RA Yoshida H., Osame M.;
 RT Cloning of rat lysosomal acid lipase cDNA and identification of the
 mutation in the rat model of Wolman's disease.";
 RL J. Lipid Res. 36:2212-2218(1995).
 CC -1- FUNCTION: CRUCIAL FOR THE INTRACELLULAR HYDROLYSIS OF CHOLESTERYL
 CC ESTERS AND TRIGLYCERIDES THAT HAVE BEEN INTERNALIZED VIA RECEPTOR-
 CC MEDIATED ENDOCYTOSIS OF LIPOPROTEIN PARTICLES. IMPORTANT IN
 CC SUPPRESSING THE EFFECT OF LDL (LOW DENSITY LIPOPROTEIN) UPTAKE ON
 CC OF ENDOGENOUS CELLULAR CHOLESTERYL ESTER FORMATION (BY
 CC similarity).
 CC -1- CATALYTIC ACTIVITY: A sterol ester + H(2)O = a sterol + a fatty
 CC acid.
 CC -1- SUBCELLULAR LOCATION: Lysosomal.
 CC -1- SIMILARITY: BELONGS TO THE ABHYDROLASE SUPERFAMILY. LIPASE FAMILY.
 CC
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@isb-sib.ch).
 CC
 CC EMBL; S81497; AAB36043.2; -;
 DR InterPro; IPR000073; A/b hydrolase.
 DR InterPro; IPR000734; Lipase.
 DR InterPro; IPR000379; Ser esters site.
 DR Pfam; PF04083; abhydro_lipase; 1.
 DR Pfam; PF00561; abhydro_lipase; 1.
 DR PROSITE; PS00120; LIPASE SER; GLYCOPROTEIN; SIGNAL; LYSOSOME.
 KW Hydrolase; Lipid degradation; Glycoprotein; Signal; Lysosome.
 FT SIGNAL 1 19
 FT CHAIN 20 397
 FT LYSOSOMAL ACID LIPASE/CHOLESTERYL ESTER
 FT HYDROLASE.
 FT ACT SITE 172 372
 FT ACT SITE 372 372
 FT CARBOHYD 34 34
 FT CARBOHYD 99 99
 FT CARBOHYD 159 159
 FT CARBOHYD 271 271
 FT CARBOHYD 319 319
 FT CARBOHYD 397 397
 FT SEQUENCE 397 AA; 45186 MW; 97A38595A0523947 CRC64;
 Query Match 54.3%; Score 1220.5; DB 1; Length 397;
 Best Local Similarity 58.2%; Pred. No. 2.7e-97;
 Matches 220; Conservative 67; Mismatches 84; Indels 7; Gaps 3;
 QY 37 PT---KAVDPEAFNMSIIHQHGPCEYEYVATEDGVLNRPGLVOPKTKGRPV 93
 DB 20 PGTGISAVDPEANMNVTEIIMWGYF--EHSVQTDGVLGVHPIPHGRKNQDPKGRPV 77
 QY 94 VLQV--HGLVGASNWIINLPNNISGLFILDAGDFVWNGSNRGNWKRKHTLSIDODEF 151
 DB 78 VYLQWREHFLADSSNWNINDDNSLGLFILDAGDFVWNGSRGNTWKRKHTLSVQDEY 137
 QY 152 WAFSDDEMAKYLPLASINILNKTQEQQLYVGVSGQGTGFIAPFSTMPQLAKIKMYFA 211
 DB 138 WAFSDDEMAKYLPLASINILNKTQEQQLYVGVSGQGTGFIAPFSTMPQLAKIKMYFA 197
 QY 212 LAPATVHKASPGKFLLLPDMIKGLGKKEELYQTRFLRLQVLYLGGVILDDQCSN 271
 DB 198 LAPVLSLIPASGPMYKGLRLLPDLLEDFGQKQFLPQSAMVKNLSTHICTHIVIKELCAN 257
 QY 272 IMLLGGFNTNNMNSRVSVAHATLACTSVQNLHWSQVNSGELRAFDSGSKNLEK 331
 DB 258 IFPLCGFNEKNLNSRVDVYTHCPAGTSVQNMVHTQVVKYHKLQAFDGGSSDKNVFH 317
 QY 332 CNOQTPVRYRVDMTVPTAMTGGQDLNSEDVQKLLSEVNTNLYHKNIPEWHAHDFIW 391
 DB 318 YNQSPYPLYSIKOMQLPALMSGGKDLADTSINILLTEIPTLVYHKNIPEWHDLOFIW 377

QY 392 GLDAPHRMNEIHLMOQ 409
 DB 378 GLDAPWRLYNEVSLMKK 395

RESULT 3

LICH MOUSE
 ID LICH_MOUSE STANDARD; PRT; 397 AA.
 AC Q9ZOM5;
 DT 28-FEB-2003 (Rel. 41, Created)
 DT 28-FEB-2003 (Rel. 41, Last sequence update)
 DT 28-FEB-2003 (Rel. 41, Last annotation update)
 DE Lysosomal acid lipase/cholesterol ester hydrolase precursor
 DE (SC 3.1.1.13) (LAL) (Acid cholesterol ester hydrolase) (Sterol
 DE esterase) (Lipase A) (Cholesteryl esterase).
 GN LIPA OR LIPI.
 OS Mus musculus (Mouse).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
 OX NCBI_TaxID=10090;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN=C57BL/6 X CBA; TISSUE=Liver;
 RX MEDLINE=96363957; PubMed=8725147;
 RA Du H., Witte D.P., Grabowski G.A.;
 RT "Tissue and cellular specific expression of murine lysosomal acid
 RT lipase mRNA and protein";
 RL J. Lipid Res. 37:937-949(1996).
 CC -1- FUNCTION: CRUCIAL FOR THE INTRACELLULAR HYDROLYSIS OF CHOLESTERYL
 CC ESTERS AND TRIGLYCERIDES THAT HAVE BEEN INTERNALIZED VIA RECEPTOR-
 CC MEDIATED ENDOCYTOSIS OF LIPOPROTEIN PARTICLES. IMPORTANT IN
 CC SUPPRESSING THE EFFECT OF LDL (LOW DENSITY LIPOPROTEIN) UPTAKE ON
 CC OF ENDOGENOUS CELLULAR CHOLESTERYL ESTER FORMATION (BY
 CC similarity).
 CC -1- CATALYTIC ACTIVITY: A sterol ester + H(2)O = a sterol + a fatty
 CC acid.
 CC -1- SUBCELLULAR LOCATION: Lysosomal.
 CC -1- TISSUE SPECIFICITY: Expressed at low levels in most tissues. High
 CC level expression is found in hepatocytes and splenic and thymic
 CC cells. Very high level expression is observed in cells of the
 CC small intestinal villi, the zona fasciculata and reticularis of
 CC the adrenal cortex, pancreatic acini, and renal tubular
 CC epithelium.
 CC -1- SIMILARITY: BELONGS TO THE ABHYDROLASE SUPERFAMILY. LIPASE FAMILY.
 CC
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@isb-sib.ch).
 CC
 CC EMBL; Z31689; CAA83494.1; -;
 DR MGD; MGI:96789; Lipi.
 DR InterPro; IPR000073; A/b hydrolase.
 DR InterPro; IPR000734; Lipase.
 DR InterPro; IPR000379; Ser esters site.
 DR Pfam; PF04083; abhydro_lipase; 1.
 DR Pfam; PF00561; abhydro_lipase; 1.
 DR PROSITE; PS00120; LIPASE SER; 1
 DR Hydrolase; Lipid degradation; Glycoprotein; Signal; Lysosome.
 KW Hydrolase; Lipid degradation; Glycoprotein; Signal; Lysosome.
 FT SIGNAL 1 19
 FT CHAIN 20 397
 FT LYSOSOMAL ACID LIPASE/CHOLESTERYL ESTER
 FT HYDROLASE.
 FT ACT SITE 172 372
 FT ACT SITE 372 372
 FT CARBOHYD 34 34
 FT CARBOHYD 99 99
 FT CARBOHYD 159 159
 FT CARBOHYD 271 271

DR	PIR; A23045; LIRTT.
DR	InterPro; IPR000073; A/b_hydrolase.
DR	InterPro; IPR000734; Lipase.
DR	InterPro; IPR000379; Ser_estra site.
DR	Pfam; PF04803; abhydro lipase; 1.
DR	Pfam; PF00561; abhydrolase; 1.
DR	PROSITE; PS00120; LIPASE SER; 1.
KW	Hydrolase; lipid degradation; Glycoprotein; Signal.
SIGNAL	1 18
FT	CHAIN 19 395 TRIACYLGLYCEROL LIPASE, LINGUAL.
FT	ACT_SITE 171 171 CHARGE RELAY SYSTEM (BY SIMILARITY).
FT	ACT_SITE 342 342 CHARGE RELAY SYSTEM (BY SIMILARITY).
FT	ACT_SITE 371 371 CHARGE RELAY SYSTEM (BY SIMILARITY).
FT	DISELFD 245 254 BY SIMILARITY.
FT	CARBOHYD 33 33 N-LINKED (GLCNAC. .) (POTENTIAL).
FT	CARBOHYD 68 68 N-LINKED (GLCNAC. .) (POTENTIAL).
FT	CARBOHYD 98 98 N-LINKED (GLCNAC. .) (POTENTIAL).
FT	CARBOHYD 270 270 N-LINKED (GLCNAC. .) (POTENTIAL).
SQ	SEQUENCE 395 AA; E60185A4923522EA CRC64;
Query Match	51.8%; Score 1164.5; DB 1; Length 395;
Best Local Similarity	54.3%; Pred. No. 1.7e-92;
Matches 217; Conservative 69; Mismatches 108; Indels 1; Gaps 1;	
QY	17 MWLLILVAVNFQ-RNVNSVMPTKAVDPFAFNISIIHQGYPCEEYEVEDGYYLSV 75 : : : : : : : : : : : : : : : :
DB	1 MWLLIITSVISTFGGAHGLFGLPGCPNPANNISQMITYNGYPCEVEVVTEGGYILGV 60
QY	76 NRIRPGIVQPCKTGSRRPVLLQHGLVGASAWISNLPNNSIGFIADAGEDPVMGNSRG 135 : : : : : : : : : : : : : : : :
DB	61 YRIIPHGNKSENGTKGPVPVYLQHGILTASNIAINLPNNSLAFLMADAGYDVVLGNSRG 120
QY	136 AWSRKHKHTLSDIDQETWFARSYDEMARPDLPAVINFIQKTQEKIYVYGYSQGTYTMGFIA 195 : : : : : : : : : : : : : : : :
DB	121 TWSRKQVYSPDSVEFWASFDEMAKYDLPATINIVQKTGOEKIHVYGHSGQTIGFIA 180
QY	196 PSTMPELAOKIKMYFALAIATVTKHAKSPTGYELLPPDMTKGLFKGKEFYIQTRFLRQL 255 : : : : : : : : : : : : : : : :
DB	181 FSTNPTLAKKIKTFFYALAPVATVKYTQSPLKKISIFITFLFKLMFGKMFLPHYTFDDFL 240
QY	256 VIVLCGVVILDQICSNIMLLLGFTNNNMGRASVVAATHLAGTSVNILHNSQAQVNSG 315 : : : : : : : : : : : : : : : :
DB	241 GTEVCSEVIDLDCSNLTIFCFGDFKNLVSRFPVILGENPAGTISVDQFLWAQLVRSG 300
QY	316 ELRAFDFGSETKXLEKCNQPTPVRYRVDRMTVPTAMWTGGQDMTJLNPDENVKMLLSVTNL 375 : : : : : : : : : : : : : : : :
DB	301 KFQAFNNGSPSQNMLHYNQKTPPEYDVSAMTVPAVAVNGNDILLADPODVAMLPLKL-SNL 360
QY	376 IYHKNIPEWARHVDFTWCILDAPHENYNEIHLMQQE 410 : : : : : : : : : : : : : : : :
DB	361 LFPHKEIILAYNHLDFTWMDAPQEVNEMISMMAED 395
RESULT 5	
LIPG CANFA	
ID	LIPG CANFA STANDARD; PRT; 398 AA.
AC	P80035; O02857;
DT	01-NOV-1991 (Rel. 20, Created)
DT	15-JUL-1998 (Rel. 36, Last sequence update)
DT	15-SEP-2003 (Rel. 42, Last annotation update)
DE	Triacylglycerol lipase, gastric precursor (EC 3.1.1.3) (Gastric
DE	lipase) (GL).
GN	LIFT.
OS	Canis familiaris (Dog).
OC	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC	Mammalia; Eutheria; Carnivora; Pissipedia; Canidae; Canis.
OX	NCBI_TaxID=9615;
RN	[1]
RP	SEQUENCE FROM N.A.
TX	TISSUE=Stomach;
RX	MEDLINE=99450174; PubMed=10520456;
RA	Vaganay S., Jolliff G., Bertaux O., Toselli E., Devignes M.D., Benicourt C.,

RT "The complete cDNA sequence encoding dog gastric lipase.";
 RL DNA Seq. 8:257-262(1998).
 RN [2]

RN SEQUENCE OF 20-59.
 RX MEDLINE=92037652; PubMed=1935982;
 RA Carriere F., Moreau H., Raphael V., Laugier R., Benicourt C.,
 RA Junien J.-L., Verger R.;
 RT "Purification and biochemical characterization of dog gastric
 RT lipase.";
 RL Eur. J. Biochem. 202:75-83(1991).
 CC -1- CATALYTIC ACTIVITY: Triacylglycerol + H(2)O = diacylglycerol + a
 CC fatty acid anion.
 CC -1- SIMILARITY: BELONGS TO THE ABHYDROLASE SUPERFAMILY. LIPASE FAMILY.
 CC -----
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@isb-sib.ch).
 CC -----

CC EMBL; Y13899; CA74198.1; -;
 CC PDB; 1X8Q; 20-MAR-02.
 DR InterPro; IPR000073; A/b hydrolase.
 DR InterPro; IPR000734; Lipase.
 DR InterPro; IPR000379; Ser_estr_1.
 DR Pfam; PF04083; abhydro_lipase; 1.
 DR Pfam; PF00561; abhydro_lipase; 1.
 DR PROSITE; PS00120; LIPASE SER; 1.
 KW Hydrolase; Lipid degradation; Glycoprotein; Signal; 3D-structure.
 FT SIGNAL 1 19
 FT CHAIN 20 398 TRIACYLGLYCEROL LIPASE, GASTRIC.
 FT ACT_SITE 172 172 CHARGE RELAY SYSTEM (BY SIMILARITY).
 FT ACT_SITE 372 372 CHARGE RELAY SYSTEM (BY SIMILARITY).
 FT CARBOHYD 34 34 N-LINKED (GLCNAC. . .) (POTENTIAL).
 FT CARBOHYD 99 99 N-LINKED (GLCNAC. . .) (POTENTIAL).
 FT CARBOHYD 271 271 N-LINKED (GLCNAC. . .) (POTENTIAL).
 FT CARBOHYD 327 327 N-LINKED (GLCNAC. . .) (POTENTIAL).
 FT CONFLICT 39 39 I -> T (IN REF. 2).
 SQ SEQUENCE 398 AA; 45130 MW; E04D62F7518E386C CRC64;

Query Match 51.4%; Score 1154; DB 1; Length 398;
 Best Local Similarity 54.3%; Pred. No. 1.4e-91;
 Matches 216; Conservative 71; Mismatches 99; Indels 12; Gaps 5;

QY 17 WLLILVAVMFORNVNSVH-----MPTKAVDPPEAFNMISLIHQGYPCBEYEVATEDG 70
 DB 1 WMLLTAASVIS-TIGTTHGLFGKLUHPT---NPEVTNMSQMTYWGYPABEEYEVATEDG 56
 QY 71 YILSVNRIIPRGVLQPKTGSPPVLLQHLGVGGASNWIISNLPNNSLGFILADAGFDVWVG 130
 DB 57 YILGIDRIPYGRKNSNIGRRPVAFLQHLGLASATNWIISNLPNNSLGFILADAGFDVWVG 116
 QY 131 NSRGNAWKKHTLISIDQDEFWAFSYDEMARFDLPVINFILQKTQBEKIYVGSQGT 190
 DB 117 NSRGNTWARNLYSPDSVEFWAFSDEWAKYDLPATIDFLKKTQDKLHYVGSQGT 176
 QY 191 MGFIAFTMPELAQIKMYFALAPATVTKHAKSPGTKEFLLPDMNKGKGFKEFLYQTR 250
 DB 177 IGFIAFTNPKLAKIKFYALAPATVTKYIETLLNKLMLVPSFLKIFGNKIF-YPHH 235
 QY 251 FLRQ-LVLYCQVILDCISNIMLLGFTNTNNMNSRASVYAAHTLAGTSVQNILHWS 309
 DB 236 FFDQFLATEVCSSRETVDLLCSNALPTICGFTDNLNMSRLDYLSHNPAGTSVQNVLHWS 295
 QY 310 QAVNSELRAPDWSGETNLEKCNQPTPVREVRDMTVPATWMTGGQDLNLPDSDVKMLL 369
 DB 296 QAVKSGKQAFDWSGSPVQNMHYSMPYPYNTLDMHVPVIAVWNGNDLLADPHDVL 355
 QY 370 SEVTNLIYHKNPPEWAHVDVFWGLDAPHRMVNIITHLM 407
 DB 356 SKLPNLIVHRKIPFPYHLDVFWAMDAPQAVNYEIVSM 393

RESULT 6

QY 18 WLLILVAVMFORNVNSVHPTKAVDPPEAFNMISLIHQGYPCBEYEVATEDG 77
 DB 3 WLLVTVCFTHMSGNAFCFLGKIAKAPASMNYSQMSIYWGYSFSEMKVITADG 62
 QY 78 IPRGLVOPKKTGSPPVLLQHLGVGGASNWIISNLPNNSLGFILADAGFDVWNGSRGNW 137
 DB 63 IPRGNANHLGQRPVVFQHLGLASATNWIISNLPNNSLGFILADAGFDVWNGSRGNW 122
 QY 138 SRKHKTLSIDQDEFWAFSYDEMARFDLPVINFILQKTQBEKIYVGSQGTMGFIAPS 197
 DB 123 AQEHLTYSDSPSEFWAFSDEWAEYDLPSTIDFLRRTQCKLHYVGSQGTGFIAPS 182
 QY 198 TMEPLAQIKMYFALAPATVTKHAKSPGTKEFLLPDMNKGKGFKEFLYQTR 256
 DB 183 TSPTLAEKIKVYFALAPATVTKYISLNFKNLAIIPHFLPKIIFGDKMP-YPHFTLEQFLG 241

Query Match 50.7%; Score 1139; DB 1; Length 397;
 Best Local Similarity 53.7%; Pred. No. 2.7e-90;
 Matches 212; Conservative 68; Mismatches 113; Indels 2; Gaps 2;

QY 18 WLLILVAVMFORNVNSVHPTKAVDPPEAFNMISLIHQGYPCBEYEVATEDG 77
 DB 3 WLLVTVCFTHMSGNAFCFLGKIAKAPASMNYSQMSIYWGYSFSEMKVITADG 62
 QY 78 IPRGLVOPKKTGSPPVLLQHLGVGGASNWIISNLPNNSLGFILADAGFDVWNGSRGNW 137
 DB 63 IPRGNANHLGQRPVVFQHLGLASATNWIISNLPNNSLGFILADAGFDVWNGSRGNW 122
 QY 138 SRKHKTLSIDQDEFWAFSYDEMARFDLPVINFILQKTQBEKIYVGSQGTMGFIAPS 197
 DB 123 AQEHLTYSDSPSEFWAFSDEWAEYDLPSTIDFLRRTQCKLHYVGSQGTGFIAPS 182
 QY 198 TMEPLAQIKMYFALAPATVTKHAKSPGTKEFLLPDMNKGKGFKEFLYQTR 256
 DB 183 TSPTLAEKIKVYFALAPATVTKYISLNFKNLAIIPHFLPKIIFGDKMP-YPHFTLEQFLG 241

QY 17 WLLILVAVMFORNVNSVH-----MPTKAVDPPEAFNMISLIHQGYPCBEYEVATEDG 70
 DB 1 WMLLTAASVIS-TIGTTHGLFGKLUHPT---NPEVTNMSQMTYWGYPABEEYEVATEDG 56
 QY 71 YILSVNRIIPRGVLQPKTGSPPVLLQHLGVGGASNWIISNLPNNSLGFILADAGFDVWVG 130
 DB 57 YILGIDRIPYGRKNSNIGRRPVAFLQHLGLASATNWIISNLPNNSLGFILADAGFDVWVG 116
 QY 131 NSRGNAWKKHTLISIDQDEFWAFSYDEMARFDLPVINFILQKTQBEKIYVGSQGT 190
 DB 117 NSRGNTWARNLYSPDSVEFWAFSDEWAKYDLPATIDFLKKTQDKLHYVGSQGT 176
 QY 191 MGFIAFTMPELAQIKMYFALAPATVTKHAKSPGTKEFLLPDMNKGKGFKEFLYQTR 250
 DB 177 IGFIAFTNPKLAKIKFYALAPATVTKYIETLLNKLMLVPSFLKIFGNKIF-YPHH 235
 QY 251 FLRQ-LVLYCQVILDCISNIMLLGFTNTNNMNSRASVYAAHTLAGTSVQNILHWS 309
 DB 236 FFDQFLATEVCSSRETVDLLCSNALPTICGFTDNLNMSRLDYLSHNPAGTSVQNVLHWS 295
 QY 310 QAVNSELRAPDWSGETNLEKCNQPTPVREVRDMTVPATWMTGGQDLNLPDSDVKMLL 369
 DB 296 QAVKSGKQAFDWSGSPVQNMHYSMPYPYNTLDMHVPVIAVWNGNDLLADPHDVL 355
 QY 370 SEVTNLIYHKNPPEWAHVDVFWGLDAPHRMVNIITHLM 407
 DB 356 SKLPNLIVHRKIPFPYHLDVFWAMDAPQAVNYEIVSM 393

GO; GO:0004806; P:triacylglycerol lipase activity; TAS.	
GO; GO:0006641; P:triacylglycerol metabolism; NAS.	
InterPro; IPR000073; A/b hydrolase.	
InterPro; IPR000734; Lipase.	
InterPro; IPR000379; Ser esters site.	
Pfam; PF04083; abhydro lipase; 1.	
Pfam; PF00561; abhydrolase; 1.	
PROSITE; PS00120; LIPASE_SER; 1.	
KW Hydrolase, Lipid degradation; Glycoprotein; Signal; Polymorphism;	
3D-structure.	
FT SIGNAL	1 19
FT CHAIN	20 398
FT ACT_SITE	172 172
FT ACT_SITE	343 343
FT ACT_SITE	372 372
FT DISULFID	246 255
FT CARBOHYD	34 34
FT CARBOHYD	99 99
FT CARBOHYD	271 271
FT CARBOHYD	327 327
FT VARIANT	161 161
FT SEQUENCE	398 AA; 45237 MW; CD3EE162IC014F0P CRC64;
Query Match	49.8%; Score 1119; DB 1; Length 398;
Best Local Similarity	52.2%; Pred. No. 1.4e-88;
Matches 210; Conservative	73; Mismatches 107; Indels 12; Gaps 2;
QY 17 MWLLILVAYMQRNVSVMPTKAV-----DPEAFWNISEIIHQGVCPEEVEVATED 69	: : : : : : : : : : : : : : : : : :
DB 1 MWLLLTWASLI-----SVLGTTHGLFGKLHPGSPEVTMNISQMITYGWPNSETEVTTD 55	: :
QY 70 GYLIVSNRI PRGLVQP KKTGRSREVVLIHQLGVGGASNINSLPNNSLGFIILADAGFDVWM 129	: : : : : : : : : : : : : : : : : : : : :
DB 56 GYLIVENRI PVGKNKGNTQRPVVFLQHGLLASATNMISLNPNNSLAIFILADAGYDVL 115	: :
QY 130 NSRGNNAWRKHKTLSIDQDEFWAFSYDEWARFDLPVINFLQTKGQEKIYYVGYSQGT 189	: : : : : : : : : : : : : : : : : : : : :
DB 116 NSRGNTWARENLVYSPDSVEFWAFSDMAKYDLPATIDFIVKTKGQKQLHYVGHSGQT 175	: :
QY 190 TWGTFARSTMDLAQIKMTFYALAPTATVKHAKSPCTFKLLPDMMWIKGLFGKEEFLYQT 249	: : : : : : : : : : : : : : : : : : : : :
DB 176 TIGTFASTNDPLAKRKITFYALAPVATVKYKSLINKLRFPVPQSLEKFIFGDKIFYPHN 235	: :
QY 250 RFLRLQVIYLCGQVILDQICSNIIMLLGLGFNTNNMMSPASVYAHTLAGTSVNILHWS 309	: : : : : : : : : : : : : : : : : : : : :
DB 236 FPDQFLATEVCSREMLNLLCSNALFITCGPSKNFNSTRSLDYLSHNPAGTSVQNMFHEWT 295	: :
QY 310 QAVNSGELRAFDMGSETNKLEKCNQTPVRYVRDMTVETAMWTGGQDWLNPNEDVKMLL 369	: : : : : : : : : : : : : : : : : : : : :
DB 296 QAVSGKGFQAIDMGSPQNRMHVDQSQPPYNYVTANVFIAVNGSKOLLADPDQDVGLLL 355	: : : : ~~~~~~ : : : : ~~~~~~ : : : : ~~~~~~ : : : : ~~~~~~ : : : :
QY 370 SEVNTNLIYHKNIPEAHVDFIWGLDAPHRYNEIHLMQOEE 411	: :
DB 356 PKGLPNLIYHKIEPFYNHLDFIWADAPQEYINDIVSMISEDK 397	: :
RESULT 8	
LIP3 DROME	
ID LIP3 DROME STANDARD; PRT; 394 AA.	
AC 046108;	
DT 16-OCT-2001 (Rel. 40, Created)	
DT 16-OCT-2001 (Rel. 40, Last sequence update)	
DT 16-OCT-2001 (Rel. 40, Last annotation update)	
DE Lipase 3 precursor (EC 3.1.1.-) (Dmlip3).	
GN Lip3 OR CG8823.	
OS Drosophila melanogaster (Fruit fly).	
OC Eukaryota; Metazoa; Arthropoda; Hexapoda; Insecta; Pterygota;	
OC Neoptera; Endopterygota; Diptera; Brachycera; Muscomorpha;	
OC Ephydroidea; Drosophilidae; Drosophila.	
OX NCBI_TaxID=7227;	
RN [1]	
RP SEQUENCE FROM N.A.	

RA Ballew R.M., Basu A., Baxendale J., Bayraktaroglu L., Beasley E.M.,
RA Beeson K.Y., Benos P.V., Berman B.P., Bhandari D., Bolshakov S.,
RA Borkova D., Botchan M.R., Bouck J., Brokstein P., Brottier P.,
RA Burtis K.C., Busan D.A., Butler H., Cadieu E., Center A., Chandra I.,
RA Cherry J.M., Cawley S., Dahke C., Davenport L.B., Davies P.,
RA de Pablos B., Delcher A., Deng Z., Mays A.D., Dew I., Dietz S.M.,
RA Dodson K., Doup L.E., Downes M., Dugan-Rocha S., Dunkov B.C., Dunn P.,
RA Durbin K.J., Evangelista C.G., Ferraz C., Ferriera S., Fleischmann W.,
RA Foster C., Gabriellian A.E., Gary N.S., Gelbart W.M., Glaser K.,
RA Glodak A., Gong F., Gorrell J.H., Gu Z., Guan P., Harris M.,
RA Harris N.L., Harvey D., Helman T.J., Hernandez J.R., Houck J.,
RA Hostin D., Houlston K.A., Howland T.J., Wei M.-H., Ibegwam C.,
RA Jalali M., Kaulush F., Karpen G.H., Ke Z., Kennison J.A., Ketchum K.A.,
RA Kimmel B.E., Kodira C.D., Kraft C., Kravitz S., Kulp D., Lai Z.,
RA Lasko P., Lei Y., Levitsky A.A., Li J., Li Z., Liang Y., Lin X.,
RA Liu X., Mattei B., McIntosh T.C., McLeod M.P., McPherson D.,
RA Merkulo G., Milshina N.V., Mobarry C., Morris J., Moshrefi A.,
RA Mount S.M., Moy M., Murphy B., Murphy L., Muzny D.M., Nelson D.L.,
RA Nelson D.R., Nelson K.A., Nixon K., Nusskern D.R., Pacleb J.M.,
RA Palazzolo M., Pittman G.S., Pan S., Pollard J., Puri V., Reese M.G.,
RA Reinert K., Remington K., Saunders R.D.C., Scheeler P., Shen H.,
RA Shue B.C., Siden-Kiamos I., Simpson M., Skupeki M.P., Smith T.,
RA Spier E., Spradling A.C., Scapleton M., Strong R., Sun E.,
RA Svirskas R., Tector C., Turner R., Venter E., Wang A.H., Wang X.,
RA Wang Z.-Y., Wasserman D.A., Weinstein G.M., Weissbach J.,
RA Williams S.M., Woodage T., Worley K.C., Wu D., Yang S., Yao Q.A.,
RA Ye J., Yeh R.-F., Zaveri J.S., Zhan M., Zhang G., Zhao Q., Zheng L.,
RA Zheng X.H., Zhong F.N., Zhong W., Zhou X., Zhu S., Zhu X., Smith H.O.,
RA Gibbs R.A., Myers E.W., Rubin G.M., Venter J.C.;
RT "The genome sequence of *Drosophila melanogaster*;"
RL Science 287:2185-2195 (2000).
[3]
RN SEQUENCE FROM N.A.
RC STRAIN=Berkeley; TISSUE=Embryo;
RX MEDLINE=22426066; PubMed=12537569;
RA George R.A., Guarin H., Krommiller B., Pacleb J.M., Park S., Wan K.H.,
RA Rubin G.M., Celniker S.E.;
RT "A *Drosophila* full-length cDNA resource;"
RL Genome Biol. 3:RESEARCH0080.1-RESEARCH0080.8 (2002).
CC -1- FUNCTION: COULD BE A DIGESTIVE ENZYME.
CC -1- SUBCELLULAR LOCATION: Secreted (Probable).
CC -1- TISSUE SPECIFICITY: In 14 hour embryos expression is seen in the
CC foregut/midgut boundary.
CC -1- DEVELOPMENTAL STAGE: Expressed from 14 hour embryos through to
CC adulthood. There is a weak maternal contribution to early embryos.
CC -1- SIMILARITY: PARTIAL WITH OTHER LIPASES (PANCREATIC, GASTRIC,
CC HEPATIC, LINGUAL, LIPOPROTEIN, BACTERIAL, ETC.). ALSO SIMILAR TO
CC LEPIDOPTERAN EG-SPECIFIC AND YOLK PROTEINS.
CC -1- CAUTION: It is uncertain whether Met-1 or Met-7 is the initiator.
CC
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation
CC of the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
CC or send an email to license@isb-sib.ch).
CC
CC EMBL; Y14366; CAA74736.1; ALT_INIT.
CC EMBL; AE003629; AAF52994.1; -
CC EMBL; AY075506; AAL68315.1; -
CC Flybase; FBN0023496; Lip1.
CC InterPro; IPR000073; A/b hydrolase.
CC InterPro; IPR000734; Lipase.
CC InterPro; IPR000379; Ser esters site.
CC Pfam; PF04083; abhydro lipase; 1.
CC Pfam; PF00561; abhydro lipase; 1.
CC PROSITE; PS00120; LIPASE_SER; FALSE_NEG.
CC HydroLase; Lipid degradation; Signal; Glycoprotein.
CC SIGNALL 1 24 POTENTIAL.
CC CHAIN 25 439 LIPASE 1.
CC DOMAIN 30 44 POLY-GLU.

FT ACT_SITE 197 197 CHARGE RELAY SYSTEM (BY SIMILARITY).
FT ACT_SITE 393 393 CHARGE RELAY SYSTEM (BY SIMILARITY).
FT CARBOHYD 124 124 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 151 151 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 346 346 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 379 379 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 426 426 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CONFLICT 10 10 I -> I (IN REF. 1).
FT CONFLICT 213 213 Y -> F (IN REF. 1).
FT CONFLICT 412 412 Q -> E (IN REF. 1).
SQ SEQUENCE 439 AA; 50660 MW; 9E32E20B8AE933F CRC64;
Query Match 24.3%; Score 546; DB 1; Length 439;
Best Local Similarity 31.8%; Pred. No. 2.8e-39;
Matches 123; Conservative 93; Mismatches 157; Indels 24; Gaps 7;
Qy 39 KAVDPEAFNMISIIHQHGPCEYEYEVATEDGYILSVNRIIPGLVQPKTKTGRPVVLLQH 98
Db 58 KNIKQDSTLSVDKLIAGYGESEVHHVTTEDGYILTMHRI-----RKQGA-PPFLLQH 109
Qy 99 GLVGGASNMISNLPNNISLGFILADAGFDVYMGNSRGNAMSRKHKTLISIDQDEFWAFSYDB 158
Db 110 GLVDSSAGFVGMGNVSLAYLLADHNDVLMGNRYSRNHTLIDPDSKFWDSWIE 169
Qy 159 MARFDLPVAVINFLQKTGQEKIYYGVSGQTTMGFTAFSTMPELAQIKKYPALAPIATV 218
Db 170 IGMVYDLPAMIDHVLKVTGPKLHYAGHSQCTSFVCMSPAYNDKVVSMQALAPAVYA 229
Qy 219 KHAKS-PGTFFLLPDMIMKLGKFKFELYQTRFLRQLVYLGGVILDDQICSNIMLLAG 277
Db 230 KETEDHPYIRALISLYNSLVGSSIREMNGEFRFLCRMT-----EETELCLIEAVFGIV 283
Qy 278 GFNTNNMNSRASVYAAHTLAGTSVQVILHWSQAVNSGELRAPDWSGSETKNLEKCNQPTP 337
Db 284 GRWNENFRKQFFVILGHYPAGVAAKQVHFQIKSGRFAPVSY-SSNKNQMLYRDHLP 342
Qy 338 VRYRVRDMVTPTAMWTGGQDNLNPDVQKMLSEVNTLYHKNIIP-EWAHVDVFIWGLDA 395
Db 343 PRYNLSLVTVTFVTYVYSTNDLLCHPKDVFESMCDLGNVTGKLVPOKEFNHMDFLWAIDV 402
Qy 396 PHMYNEIIHLM-----QOEETNLSQ 416
Db 403 RKMLYRRMLQVLGKVPESPEANRNR 429
RESULT 10
TGL1 YEAST
ID TGL1 YEAST STANDARD; PRT; 548 AA.
AC P34163;
DT 01-FEB-1994 (Rel. 28, Created)
DT 01-FEB-1994 (Rel. 28, Last sequence update)
DT 28-FEB-2003 (Rel. 41, Last annotation update)
DE Triglyceride lipase-cholesterol esterase (EC 3.1.1.-).
GN TGL1 OR YKL140W OR YKL5.
OS Saccharomyces cerevisiae (Baker's yeast).
OC Eukaryota; Fungi; Ascomycota; Saccharomycotina; Saccharomycetes;
OC Saccharomycetales; Saccharomycetaceae; Saccharomycetes.
OX NCBI_TaxID=4932;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=92245761; PubMed=1574929;
RA Abraham P.R., Mulder A., Van't Riet J., Planta R.J., Raue H.A.;
RT "Molecular cloning and physical analysis of an 8.2 kb segment of
RT chromosome XI of *Saccharomyces cerevisiae* reveals five tightly linked
RT genes.";
RL Yeast 8:227-238 (1992).
RN [2]
RP SEQUENCE FROM N.A.
RA Rad M.R., Xu G., Kirchath L., Fritz C., Keuchel H., Hollenberg C.P.;
RL Submitted (MAR-1994) to the EMBL/GenBank/DBJ databases.
CC -1- SIMILARITY: Belongs to the AB hydrolase superfamily.
CC
CC This SWISS-PROT entry is copyright. It is produced through a collaboration

R.A. Overbeek R., Kirkness E.F., Weinstein K.G., Merrick J.M., Glodek A.,
Poulsen C.E., Adams M.D., Reich C.L.,
Kirkness E.F., Weinstein K.G., Merrick J.M., Glodek A.,
Overbeek R., Kirkness E.F., Weinstein K.G., Merrick J.M., Glodek A.,

GN ORC-2 OR F59E10.1.

GN	ORC-2
OS	Caenor

GN ORC-2 OR F59E10.1.
OS Caenorhabditis elegans.

GN ORC-2 OR P59E10.1.
OS *Caenorhabditis elegans*.

GN ORC-2 OR F59E10.1.
OS Caenorhabditis elegans.

GN ORC-2 OR F59E10.1.
OS *Caenorhabditis elegans*.

GN ORC-2 OR F59E10.1.
OS *Caenorhabditis elegans*.

GN ORC-2 OR F59E10.1.
OS *Caenorhabditis elegans*.

GN ORC-2 OR F59E10.1.
OS Caenorhabditis elegans.

GN ORC-2 OR F59E10.1.
OS *Caenorhabditis elegans*.

GN ORC-2 OR F59E10.1.
OS *Caenorhabditis elegans*.

GN ORC-2 OR F59E10.1.
OS *Caenorhabditis elegans*.

OC Eukaryota; Metazoa; Nematoda; Chromadorea; Rhabditida; Rhabditoidea;
 OC Rhabditidae; Peloderinae; Caenorhabditis.
 OX NCBI_TaxID=6239;
 RN [1]
 RN SEQUENCE FROM N.A.
 RX MEDLINE=96099401; PubMed=7502077;
 RA Gavin K.A., Hidaka M., Stillman B.;
 RT "Conserved initiator proteins in eukaryotes";
 RL Science 270:1667-1671(1995).
 RN [2]
 RN SEQUENCE FROM N.A.
 RP STRAIN=Bristol N2;
 RA Swinburne J.;
 RN Submitted (AUG-1994) to the EMBL/GenBank/DBJ databases.
 RN [3]
 RN REVISIONS.
 RP Jones S.J.M.;
 RA Submitted (MAR-1997) to the EMBL/GenBank/DBJ databases.
 CC -1- FUNCTION: COMPONENT OF THE ORIGIN RECOGNITION COMPLEX (ORC) THAT
 CC BINDS ORIGIN OF REPLICATION. IT HAS A ROLE IN BOTH CHROMOSOMAL
 CC REPLICATION AND MATING TYPE TRANSCRIPTIONAL SILENCING. BINDS TO
 CC THE ARS CONSENSUS SEQUENCE (ACS) OF ORIGINS OF REPLICATION IN AN
 CC ATP-DEPENDENT MANNER.
 CC -1- SUBUNIT: ORC IS COMPOSED OF SIX SUBUNITS (BY SIMILARITY).
 CC -1- SUBCELLULAR LOCATION: Nuclear.
 CC -1- SIMILARITY: BELONGS TO THE ORC2 FAMILY.
 CC -----
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@isb-sib.ch).
 CC -----
 DR EMBL; U40270; AAC48954.1; -;
 DR EMBL; Z36949; CAA85415.1; -;
 DR PIR; T23001; T23001.
 DR WormPep; F59E10.1; CE11492.
 DR Pfam; PF04084; ORC2; 1.
 DR DNA replication, Nuclear protein.
 SQ SEQUENCE 430 AA; 49319 MW; 54FC086BB4AD9670 CRC64;

Query Match 4.7%; Score 105; DB 1; Length 430;
 Best Local Similarity 21.9%; Pred. No. 0.21;
 Matches 61; Conservative 42; Mismatches 87; Indels 88; Gaps 13;
 QY 85 PKKTGRPVVLLQHLVGGASNLSLPPN-----SLGP-----ILADAGFDVW--MGN 131
 Db 25 PEKEGR-----QKTKNGKNSRNLQSNLEEDLEQLGFEDTSTVMAQSAIENYFMQK 78
 QY 132 S-----RGNWSRKH-----TUSIDODEFWAFSYDEMARFDLPVINFILQKTQEKTY 181
 Db 79 SASERMNNNAKSRRRAGNAGNTEIEEDDEISNAITDFTKCDLPGLRNVTIKKDNTEFEK 138
 QY 182 YGVGSQGTWTFIAFTWPELAKQIKVYFALAPIATVVKHAKSGTKELLLPDMIKGLFG 241
 Db 139 RLEHLANDFG-----KWKLYLAG-----FNLLHGVS 168
 QY 242 KKEFLYQTRFLRQLRVILCGVILDQICSNMILILGGFNTN---NMMSRASVYAAHTLA 298
 Db 169 KRDLV--TEFENELSDTYMEVDARKDGLNVKVLGAINENMKLNCNKEG-----217
 QY 299 GTSVQNLHWSQAV-----NSGEL-----RPFWSGE 325
 Db 218 ----QSTISWARSIRRRKNSQQLTLLIDNTEAPDWRSD 251

RESULT 13

ID PAPA_CAVPO STANDARD; PRT; 436 AA.

AC P70683;

DT 01-NOV-1997 (Rel. 35, Created)

DT 01-NOV-1997 (Rel. 35, Last sequence update)
 DT 28-FEB-2003 (Rel. 41, Last annotation update)
 DE Platelet-activating factor acetylhydrolase precursor (EC 3.1.1.47)
 DE (PAF acetylhydrolase) (PAF 2-acetylhydrolase) (IDL-associated
 DE phospholipase A2) (IDL-PLA(2)) (2-acetyl-1-alkylglycerophosphocholine
 DE esterase) (1-alkyl-2-acetyl-1-alkylglycerophosphocholine esterase).
 GN PLA2G7 OR PAFAH
 OS Cavia porcellus (Guinea pig).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Mammalia; Eutheria; Rodentia; Hystricognathi; Caviidae; Cavia.
 OX NCBI_TaxID=10141;
 RN [1]
 RN SEQUENCE FROM N.A.
 RC STRAIN=Hartley; TISSUE=Liver;
 RX MEDLINE=97103479; PubMed=8947850;
 RA Karasawa K., Kuge O., Kawasaki K., Nishijima M., Nakano Y., Tomita M.,
 RT Yokoyama K., Setaka M., Nojima S.;
 RT "Cloning, expression and characterization of plasma
 RT platelet-activating factor-acetylhydrolase from guinea pig.";
 RL J. Biochem. 120:838-844(1996).
 CC -1- FUNCTION: MODULATES THE ACTION OF PLATELET-ACTIVATING FACTOR (PAF)
 CC BY HYDROLYZING THE SN-2 ESTER BOND TO YIELD THE BIOLOGICALLY
 CC INACTIVE LISO-PAF. HAS A SPECIFICITY FOR SUBSTRATES WITH A SHORT
 CC RESIDUE AT THE SN-2 POSITION. IT IS INACTIVE AGAINST LONG-CHAIN
 CC PHOSPHOLIPIDS.
 CC -1- CATALYTIC ACTIVITY: 2-acetyl-1-alkyl-sn-glycero-3-phosphocholine +
 CC H2O = 1-alkyl-sn-glycero-3-phosphocholine + acetate.
 CC -1- SUBCELLULAR LOCATION: Extracellular.
 CC -1- TISSUE SPECIFICITY: Plasma.
 CC -1- SIMILARITY: BELONGS TO THE ABHYDROLASE SUPERFAMILY. LIPASE FAMILY.
 CC -----
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@isb-sib.ch).
 CC -----
 DR EMBL; D67037; BAA11054.1; -;
 DR PIR; JC5021; JC5021.
 DR InterPro; IPR000734; Lipase.
 DR InterPro; IPR005065; PAF-AH_P.II.
 DR InterPro; IPR000379; Serestrs_site.
 DR Pfam; PF03403; PAF-AH_P.II; 1.
 DR PROSITE; PS00120; LIPASE_SER; 1.
 KW Hydrolase; Lipid degradation; Glycoprotein; Signal.
 FT SIGNAL 1 21
 FT CHAIN 22 436
 FT PLATELET-ACTIVATING FACTOR
 FT ACETYLHYDROLASE
 FT ACT SITE 271 271
 FT ACT SITE 294 294
 FT ACT SITE 349 349
 FT CARBOHYD 76 76
 FT CARBOHYD 200 200
 FT CARBOHYD 324 324
 FT SEQUENCE 436 AA; 49062 MW; C359D96E392FEE11 CRC64;

Query Match 4.5%; Score 101; DB 1; Length 436;
 Best Local Similarity 20.4%; Pred. No. 0.48;
 Matches 85; Conservative 60; Mismatches 149; Indels 122; Gaps 21;
 QY 21 ILVAMFQFNVNVMPTKAVD--PEAFNITSEIHOQGYCEEV-EVATED-----GY 71
 Db 70 LMSGYTNQSSFLRLYPYSDNDFFDALW-----IPNEYFQGLTETLGLASSFLGK 119
 QY 72 ILVS---NRIPRGLVOPKKTGSR-PVYLQHLVGGASNLSLPPNLSGLFILLADAGFD 126
 Db 120 LLKLLYGVSKVPKWNPLKTKGKYLPIIPSHGL--GAFTSI-----YSAIGIELASHGFI 173
 QY 127 V-----WNGNRRGNWSRKHKHTLSIDODE-----FW 152
 Db 174 VAAVEHRDESAATYFQDAPAESGNRSMTIYKVGNLTEERKQRLQRGECSQALSW 233

QY 153 AFSYDEMARPDPVAVINFILOK-----TGOEKIYYVYSGQITMGFIATSTMPDLAQIKM 208
 Db 234 LUSIDEGEVKXVLDNFIQKGLSLDRSKVAIIHSHFG---GATVITQTLSE-DQRFRC 289
 QY 209 YFAL-----APIATVHXAKSPGKFKLLPDMIMKGLFGKK-BFLYQTRFLRLQV----- 256
 Db 290 GIALDPMPFVGVGDEYHVKIPQLPFINSEYFOSANDTKIEKFYQPKERKXIAVKGSVH 349
 QY 257 -----IYLCGGVILDOICSNIMLLGGFNNTNMNM-----SRASVAAHTLAGTSVQNLH 307
 Db 350 HNFVDFTFATGKII-----GQMLSLKGIKIDSEVAMDINKASLAFLOKYLGLD-KNFDQ 402
 QY 308 WSOAVNSGELRAFWDGSETKNLEKCNQPTFVRYRVRDMTPTAMTGGQDWLSNPE 363
 Db 403 WNSLME-----GDDENL-----IPEITPTMQSSGTGTQRNPD 436

RESULT 14

Y193 HAEIN
 ID Y193 HAEIN STANDARD; PRT; 287 AA.
 AC Q57427; O05013;
 DT 01-NOV-1997 (Rel. 35, Created)
 DT 01-NOV-1997 (Rel. 35, Last sequence update)
 DT 28-FEB-2003 (Rel. 41, Last annotation update)
 DE Putative esterase/libase HI0193 (EC 3.1.1.-).
 GN HI0193.
 OS Haemophilus influenzae.
 OC Bacteria; Proteobacteria; Gammaproteobacteria; Pasteurellales;
 OC Pasteurellaceae; Haemophilus.
 OX NCBI_TaxID=727;
 RN [1]

SEQUENCE FROM N.A.
 RC STRAIN=RD / KW20 / ATCC 51907;
 RX MEDLINE=95350630; PubMed=7542800;
 RA Fleischmann R.D., Adams M.D., White O., Clayton R.A., Kirkness E.F.,
 RA Kerlavage A.R., Bult C.J., Tomb J.-F., Dougherty B.A., Merrick J.M.,
 RA McKenney K., Sutton G., Fitzhugh W., Fields C.A., Gocayne J.D.,
 RA Scott J.D., Shirley R., Liu L.-I., Glodek A., Kelley J.M.,
 RA Weidman J.F., Phillips C.A., Spriggs T., Hedblom E., Cotton M.D.,
 RA Utterback T.R., Hanna M.C., Nguyen D.T., Studek D.M., Brandon R.C.,
 RA Fine L.D., Fritchman J.L., Fuhmann J.L., Geoghegan N.S.M.,
 RA Gnehm C.L., McDonald L.A., Small K.V., Fraser C.M., Smith H.O.,
 RA Venter J.C.;
 RT "Whole-genome random sequencing and assembly of Haemophilus influenzae
 Rd.";
 RL Science 269:496-512(1995).
 CC -!- SIMILARITY: BELONGS TO A FAMILY OF ESTERASES THAT GROUPS TOGETHER
 CC PSEUDOMONA TROPINESTERASE, DMPD, TODF AND XYL.
 CC -----

CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@isb-sib.ch).
 CC -----

DR EMBL; U32704; AAC21862.1; -
 DR PIR; E64053; E64053.
 DR TIGR; HI0193; -
 DR InterPro; IPR000073; A/b hydrolase.
 DR InterPro; IPR003089; AB_Hydrolase.
 DR InterPro; IPR000379; Ser_estrases.
 DR Pfam; PF00561; abhydrolase; 1.
 DR PRINTS; PRO011; ABHYDROLASE.
 DR KW Hypothetical protein; Hydrolase; Serine esterase; Complete proteome.
 FT ACT SITE 119 BY SIMILARITY.
 FT ACT SITE 266 BY SIMILARITY.
 SQ SEQUENCE 287 AA; 32987 MW; F28548619C060619 CRC64;

Query Match 4.3%; Score 96.5; DB 1; Length 287;
 Best Local Similarity 19.1%; Pred. No. 0.67;

Matches 62; Conservative 39; Mismatches 119; Indels 105; Gaps 12;
 QY 84 QPKKTSRPVVLQHLGVGASNWNLSNLSGLFLADAGDFVMNGRGNWASRKHKT 143
 Db 39 QVQQTINTPVLITFHLGLGDMON-----LGVI-----ARAFSEHYST 75
 QY 144 LSIDQDEFWAFSDEMARFDLPV-INTILOKQGEKIYYVYSGQITMGFIATSTMPDL 202
 Db 76 LRIDLNRHGHGSFSEKKNYQLMAEDVIAVRHLNLSKVILGHSMGCKTAMKITALCPDL 135
 QY 203 AQIKMYFALLAPIATVYKHAQSPOTKFLLPDM-----MTKGLFGKGFYLYQTR 250
 Db 136 VEK-----LIVDMSPEYEGFHKDVFNGFLFAVKNAPEN- 171
 QY 251 FLRLVLYLGGVILDOICSNIMLLGGFNNTNMNRASVAAHTLAGTSVQNLHWSQ 310
 Db 172 --RQAKPILKQEIENDE--DVQFMUKSFDVNSADCFRNL-----TALFNNTANIMDWE- 222
 QY 311 AVNSGELRAFWDGSETKNLEKCNQPTFVRYRVRDMTPTAMTGGQDWLSNPEVDVQLLS 370
 Db 223 -----KVRVF-----TPTLFIKGSSVYKIENSEKILE 251
 QY 371 EVTNLIYH--KNIPFWAHV---DPI 390
 Db 252 QFFNATAFTINGSCHVWFAKPDV 276

RESULT 15

L1N1_HUMAN
 ID L1N1_HUMAN STANDARD; PRT; 1259 AA.
 AC P08547;
 DT 01-AUG-1988 (Rel. 08, Created)
 DT 01-AUG-1988 (Rel. 08, Last sequence update)
 DT 01-AUG-1988 (Rel. 08, Last annotation update)
 DE LINE-1 reverse transcriptase homolog.
 OS Homo sapiens (Human).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
 OX NCBI_TaxID=9606;
 RN [1]
 RP SEQUENCE.
 RX MEDLINE=86230917; PubMed=2423883;
 RA Hattori M., Kuhara S., Takenaka O., Sakaki Y.;
 RT "Li family of repetitive DNA sequences in primates may be derived
 RT from a sequence encoding a reverse transcriptase-related protein.";
 RL Nature 321:625-628(1986).
 CC -!- MISCELLANEOUS: THIS SEQUENCE WAS CONSTRUCTED FROM AN ALIGNMENT OF
 CC PUBLISHED AND UNPUBLISHED SEQUENCES, DETERMINED IN VARIOUS
 CC LABORATORIES, BELONGING TO THE LINE-1 FAMILY.
 DR PIR; A25313; GHULL.
 DR InterPro; IPR005135; Exo_endo_phos.
 DR InterPro; IPR000477; RVTse.
 DR Pfam; PF03372; Exo_endo_phos; 1.
 DR Pfam; PF00078; rvt; 1.
 KW RNA-directed DNA polymerase.
 SQ SEQUENCE 1259 AA; 147112 MW; F3BC18A3803919F5 CRC64;

Query Match 4.3%; Score 95.5; DB 1; Length 1259;
 Best Local Similarity 19.7%; Pred. No. 5.9;
 Matches 90; Conservative 67; Mismatches 170; Indels 129; Gaps 23;

QY 28 QRNVNSVHMPTKAVDPDEAFNMI---SEIIHQGYPCPEEY----- 64
 Db 442 QEEVESLNRPITSSEIEALINSPLNPKSPQEPGEFTAEFYQRYKEELVPFLKLQSIKKE 501
 QY 65 -VATEDGYILSVNRIPR-GLVQPKTKGSRPVL-----LQHLVGGASNWSNL-PNN 114
 Db 502 GILPNSFYEAISIIIPKGRDRTTKENFRPISLWNIDAKILNKILANQIQOHKIKLIHD 561
 QY 115 SLGFIADAGDFVMNGSR-----GNASRKHKTLSIDQDEFWAFSYDEM----- 159
 Db 562 QVGFIPAMQG---WFNIRKSINIIQHNRTKOTNHEIISDAEK-----AFDKIQOFFMLK 614

```

Qy 160 -----ARFDPVAVINFILOKTOEKIYYVGVSGTMTGFIASFMPELA 203
Db 615 PLNKLGDGYLYKIRAIYDKPTA-NIIL--NGQ-KLEAPPLKTGTROGCPSPLEPNIV 670
Qy 204 QKIMWYFAPALAPIATVVKHAKSPGTFKLLPDMIMIKGLFGKKEFLYQTRFLRQLVIVYLCGOV 263
Db 671 LEV-----LA--RAIROKE-----IKGIQLGKEEVKLSLFADDMIVYLENFI 711
Qy 264 ILDOICSNIMLLGGFN-----TNNMMSPASVYAAHTLAGTSVQNIHWSQAVNSGELR 318
Db 712 VSAQ---NLLKLSNFSKVSQYKINVOKSOAFLYTNNROTSEQIMSELPFTTIASKRIKYL 768
Qy 319 AFDWGSSEKLEKCNQPTPVRYVRDWT-----VPTAMWTGGQDMLSNPEDVKMLLSEVT 373
Db 769 GIOLTRDVKDLFKENY-KPLNEIKEDTNNKKNIPCS-----WVGRINIVKMAI--LP 818
Qy 374 NLIYHKN-IP-----EWAHVDFIWLGDAPH 397
Db 819 KVIYRFNAIPKLPMTFFTELEKTTLKFIVNOKRAH 854

```

Search completed: February 19, 2004, 14:21:11
Job time : 14.8856 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2004 Compugen Ltd.

OM protein - protein search, using sw model

Run on: February 19, 2004, 14:17:14 ; Search time 30.1771 Seconds
(without alignments)
3617.186 Million cell updates/sec

Title: US-10-042-431-47

Perfect score: 2247

Sequence: 1 MLETLRQWIVSRHMEKMLL.....IHLMQEETNLQGRCEAVL 423

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 830525 seqs, 258052604 residues

Total number of hits satisfying chosen parameters: 830525

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database :

SPTREMBL_23.*

1: sp_archaea.*

2: sp_bacteria.*

3: sp_fungi.*

4: sp_human.*

5: sp_invertebrate.*

6: sp_mammal.*

7: sp_mhc.*

8: sp_organelle.*

9: sp_phase.*

10: sp_plant.*

11: sp_rodent.*

12: sp_virus.*

13: sp_vertebrate.*

14: sp_unclassified.*

15: sp_rvirus.*

16: sp_bacteriap.*

17: sp_archaeap.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	1991	88.6	374	4	Q96LG2
2	1975.5	87.9	422	11	Q8K2A6
3	1971.5	87.7	422	11	Q8BJ28
4	1222.5	54.4	397	11	Q8C2G7
5	1134.5	50.5	395	11	Q9D798
6	1134.5	50.5	395	11	Q9D606
7	1134.5	50.5	395	11	Q9D6P3
8	1134.5	50.5	395	11	Q9D6X0
9	1133.5	50.4	395	11	Q9CP77
10	1133.5	50.4	395	11	Q9CP88
11	1131.5	50.4	395	11	Q9D6L9
12	1129.5	50.3	395	11	Q9D6F5
13	1128.5	50.2	395	11	Q9D767
14	1127.5	50.2	395	11	Q9D7C5
15	1126.5	50.1	395	11	Q9D6L1
16	1126.5	50.1	395	11	Q9D766

17	1125.5	50.1	395	11	Q9D6Q3	Q9d6q3 mus musculu
18	1124.5	50.0	395	11	Q9D760	Q9d760 mus musculu
19	1124.5	50.0	395	11	Q9D6S5	Q9d6s5 mus musculu
20	1122.5	50.0	395	11	Q9D796	Q9d796 mus musculu
21	1119.5	49.8	395	11	Q9D6N8	Q9d6n8 mus musculu
22	1116.5	49.7	395	11	Q9D6Q2	Q9d6q2 mus musculu
23	1115.5	49.6	395	11	Q9D6M9	Q9d6m9 mus musculu
24	1098	48.9	398	11	Q8BM14	Q8bm14 mus musculu
25	829.5	36.9	275	13	Q8AYB8	Q8ayb8 pseudopleur
26	755.5	33.6	684	5	Q95XV1	Q95xv1 caenorhabdi
27	752.5	33.5	603	5	Q61866	Q61866 caenorhabdi
28	751	33.4	411	5	Q20449	Q20449 caenorhabdi
29	730	32.5	405	5	Q93789	Q93789 caenorhabdi
30	716	31.9	404	5	Q14956	Q14956 caenorhabdi
31	692	30.8	411	5	Q94252	Q94252 caenorhabdi
32	668.5	29.8	351	5	Q95X33	Q95x33 caenorhabdi
33	666	29.6	422	5	Q9VKT8	Q9vkt8 drosophila
34	655.5	29.2	426	5	Q17766	Q17766 caenorhabdi
35	655.5	29.2	434	5	Q9VKT9	Q9vkt9 drosophila
36	650.5	28.9	531	5	Q9VKS5	Q9vks5 drosophila
37	636	28.3	398	5	Q9V796	Q9v796 drosophila
38	625.5	27.8	616	5	Q77107	Q77107 piodia inte
39	602	26.8	429	5	Q8MWX1	Q8mwx1 dictyosteli
40	586.5	26.1	421	5	Q9VBK6	Q9vbk6 drosophila
41	574.5	25.6	559	5	Q17219	Q17219 bombyx mori
42	567.5	25.3	457	5	Q9VKT2	Q9vkt2 drosophila
43	567.5	25.3	457	5	Q8T3X7	Q8t3x7 drosophila
44	559	24.9	399	5	Q9VPE9	Q9vpe9 drosophila
45	535	23.8	416	5	Q9VQQ5	Q9vqq5 drosophila

ALIGNMENTS

RESULT 1

Q96LG2	PRELIMINARY;	PRT;	374 AA.
ID	Q96LG2		
AC	Q96LG2;		
DT	01-DEC-2001 (TrEMBLrel. 19, Created)		
DT	01-DEC-2001 (TrEMBLrel. 19, Last sequence update)		
DT	01-MAR-2003 (TrEMBLrel. 23, Last annotation update)		
DE	BA30415.1 (Novel lipase) (Fragment).		
GN	BA30415.1.		
OS	Homo sapiens (Human).		
OC	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;		
OC	Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.		
OX	NCBI_TaxID=9606;		
RN	[1]		
RP	SEQUENCE FROM N.A.		
RA	Bray-Allen S.;		
RL	Submitted (AUG-2001) to the EMBL/GenBank/DBJ databases.		
DR	EMBL; AL353113; CAC78754.1; -		
DR	InterPro; IPR000073; A/b hydrolase.		
DR	InterPro; IPR006693; abhydro_lipase.		
DR	InterPro; IPR000734; Lipase.		
DR	InterPro; IPR000379; Ser esters_site.		
DR	Pfam; PF00561; abhydro_lipase; 1.		
DR	Pfam; PF04083; abhydro_lipase; 1.		
DR	PROSITE; PS00120; LIPASE_SER; 1.		
FT	NON TER		
SQ	SEQUENCE 374 AA; 42371 MW; 5E7220A889437337 CRC64;		

Query Match 88.6%; Score 1991; DB 4; Length 374;
Best Local Similarity 100.0%; Pred. No. 8.4e-174;
Matches 374; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy	50	SEIIHQGVPCSEYEVATEDGYILSVNRIIPRGIVOPKTKGSRPVLLQHLGVGGASNMIS 109	
Db	1	SEIIHQGVPCSEYEVATEDGYILSVNRIIPRGIVOPKTKGSRPVLLQHLGVGGASNMIS 60	
Qy	110	NLPNNISGLIFADAGFDVWGNISRGNAWGRKHKTLISIDQDEFWAFSYDEMAFDFLPAVIN 169	
Db	61	NLPNNISGLIFADAGFDVWGNISRGNAWGRKHKTLISIDQDEFWAFSYDEMAFDFLPAVIN 120	


```

Qy 170 FILQKQZQEKIYVGVSGTGTGFIAPSTMPPELAQIKMYFALAPIATVTKHAKSPGTGKPL 229
Db 121 FILQKQZQEKIYVGVSGTGTGFIAPSTMPPELAQIKMYFALAPIATVTKHAKSPGTGKPL 180
Qy 230 LLPDMMIKGLFGKKEFLYQTRFLRQLVYLGGQVILDOICSNIMLLGGFNNTNNMSRA 289
Db 181 LLPDMMIKGLFGKKEFLYQTRFLRQLVYLGGQVILDOICSNIMLLGGFNNTNNMSRA 240
Qy 290 SVYAAHTLAGTSVONILHWSQAVNSGELRAFDWGSSEKCNQKQPTPVRYVRDVTPT 349
Db 241 SVYAAHTLAGTSVONILHWSQAVNSGELRAFDWGSSEKCNQKQPTPVRYVRDVTPT 300
Qy 350 AMWTGGQDWSNEDVKKLLSEVNTLIYHKNIPWAHVDFIWLGDAPHERMYNEIHLMOQ 409
Db 301 AMWTGGQDWSNEDVKKLLSEVNTLIYHKNIPWAHVDFIWLGDAPHERMYNEIHLMOQ 360
Qy 410 EETNLSQRCRAVL 423
Db 361 EETNLSQRCRAVL 374

```

RESULT 2

```

Q8K2A6 Q8K2A6 PRELIMINARY; PRT; 422 AA.
AC Q8K2A6;
DT 01-OCT-2002 (TremBrel. 22, Created)
DT 01-OCT-2002 (TremBrel. 22, Last sequence update)
DT 01-MAR-2003 (TremBrel. 23, Last annotation update)
DE RIKEN CDNA 453427C23 gene (BA30415.1).
OS Mus musculus (Mouse).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
OX NCBI_TaxID=10090;
RN [1]
RP SEQUENCE FROM N.A.
RA Strausberg R.;
RL Submitted (JUN-2002) to the EMBL/GenBank/DBJ databases.
[2]
RP SEQUENCE FROM N.A.
RC STRAIN=C57BL/6J; TISSUE=Mammary gland, and Vagina;
RX MEDLINE=22354683; PubMed=12466851;
RA The FANTOM Consortium,
RA the RIKEN Genome Exploration Research Group Phase I & II Team;
RT "Analysis of the mouse transcriptome based on functional annotation of
RT 60,770 full-length cDNAs."
RL Nature 420:563-573(2002).
DR EMBL; BC031933; AAH31933.1; -
DR EMBL; AK037091; BAC29699.1; -
DR InterPro; IPR000073; A/b hydrolase.
DR InterPro; IPR006693; abhydro_lipase.
DR InterPro; IPR000734; Lipase.
DR InterPro; IPR000379; Ser estsr_site.
DR Pfam; PF00561; abhydro_lipase; 1.
DR Pfam; PF04083; abhydro_lipase; 1.
DR PROSITE; PS00120; Lipase_SER; 1.
SQ SEQUENCE 422 AA; 48253 MW; C6EFBC140963E051 CRC64;

```

```

Query Match 87.9%; Score 1975.5; DB 11; Length 422;
Best Local Similarity 87.2%; Pred. No. 2.7e-172;
Matches 369; Conservative 23; Mismatches 30; Indels 1; Gaps 1;

Qy 1 MLETLRQWIVSHRMEMWLLILVAYMFQNNVSVHMTKAVDPEAFNISEIHOQGYPC 60
Db 1 MSELRSRVMTVSHRVEIWLILVAYLLQNNVNSGHLPTKAAADPEAFNVSEIHKGYPS 60
Qy 61 EYEYVATEDGYLSVNRIPRGLVOPKKTGSRPVVLLQHLGVGGASWISNLPNNSLGFIL 120
Db 61 EYEYVATEDGYLSVNRIPRQTRLKEGSRPVVLLQHLGVGGASWISNLPNNSLGFIL 120
Qy 121 ADAGFDVWNGSRGNAMSRKHTLSIDQDEFWAFSYDEMAFDFLPAVINFILOKTGOEKI 180

```

```

Db 121 ADAGFDVWNGSRGNAMSRKHTLSIDQDEFWAFSYDEMAFDFLPAVINFILOKTGOEKI 180
Qy 181 YVGVSGTGTGFIAPSTMPPELAQIKMYFALAPIATVTKHAKSPGTGKPL 240
Db 181 YVGVSGTGTGFIAPSTMPPELAQIKMYFALAPIATVTKHAKSPGTGKPL 240
Qy 241 GKKEFLYQTRFLRQLVYLGGQVILDOICSNIMLLGGFNNTNNMSRAVYAAHTLAGT 300
Db 241 GRQEFLYQTRFRQLFYILCGQVILDOICSNIIILLGGFNNTNNMSRANVYVHTPAGT 300
Qy 301 SVQNILHWSQAVNSGELRAFDWGSSEKCNQKQPTPVRYVRDVTPTAMWTGGQDWSL 360
Db 301 SVQNILHWSQAVNSGELRAFDWGSSEKCNQKQPTPVRYVRDVTPTAMWTGGQDWSL 360
Qy 361 NPEDVKLLSEVNTLIYHKNIPWAHVDFIWLGDAPHERMYNEIHLMOQSETNLSQRCR 420
Db 361 NPEDVKLLSEVNTLIYHKNIPWAHVDFIWLGDAPHERMYNEIHLMOQSETNLSQRCR 419
Qy 421 AVL 423
Db 420 VKL 422

RESULT 3
Q8BJ28 Q8BJ28 PRELIMINARY; PRT; 422 AA.
AC Q8BJ28;
DT 01-MAR-2003 (TremBrel. 23, Created)
DT 01-MAR-2003 (TremBrel. 23, Last sequence update)
DT 01-MAR-2003 (TremBrel. 23, Last annotation update)
DE BA30415.1.
OS Mus musculus (Mouse).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
OX NCBI_TaxID=10090;
RN [1]
RP SEQUENCE FROM N.A.
RA STRAIN=C57BL/6J; TISSUE=Skin;
RX MEDLINE=22354683; PubMed=12466851;
RA The FANTOM Consortium,
RA the RIKEN Genome Exploration Research Group Phase I & II Team;
RT "Analysis of the mouse transcriptome based on functional annotation of
RT 60,770 full-length cDNAs."
RL Nature 420:563-573(2002).
DR EMBL; AK037214; BAC29757.1; -
DR SEQUENCE 422 AA; 48287 MW; 5F2522D0E27050C8 CRC64;

Query Match 87.7%; Score 1971.5; DB 11; Length 422;
Best Local Similarity 87.0%; Pred. No. 6.2e-172;
Matches 368; Conservative 23; Mismatches 31; Indels 1; Gaps 1;

Qy 1 MLETLRQWIVSHRMEMWLLILVAYMFQNNVSVHMTKAVDPEAFNISEIHOQGYPC 60
Db 1 MSELRSRVMTVSHRVEIWLILVAYLLQNNVNSGHLPTKAAADPEAFNVSEIHKGYPS 60
Qy 61 EYEYVATEDGYLSVNRIPRGLVOPKKTGSRPVVLLQHLGVGGASWISNLPNNSLGFIL 120
Db 61 EYEYVATEDGYLSVNRIPRQTRLKEGSRPVVLLQHLGVGGASWISNLPNNSLGFIL 120
Qy 121 ADAGFDVWNGSRGNAMSRKHTLSIDQDEFWAFSYDEMAFDFLPAVINFILOKTGOEKI 180
Db 121 ADAGFDVWNGSRGNAMSRKHTLSIDQDEFWAFSYDEMAFDFLPAVINFILOKTGOEKI 180
Qy 181 YVGVSGTGTGFIAPSTMPPELAQIKMYFALAPIATVTKHAKSPGTGKPL 240
Db 181 YVGVSGTGTGFIAPSTMPPELAQIKMYFALAPIATVTKHAKSPGTGKPL 240
Qy 241 GKKEFLYQTRFLRQLVYLGGQVILDOICSNIMLLGGFNNTNNMSRAVYAAHTLAGT 300
Db 241 GRQEFLYQTRFRQLFYILCGQVILDOICSNIIILLGGFNNTNNMSRANVYVHTPAGT 300
Qy 301 SVQNILHWSQAVNSGELRAFDWGSSEKCNQKQPTPVRYVRDVTPTAMWTGGQDWSL 360

```



```

RESULT 6
Q9D6Q6 PRELIMINARY; PRT; 395 AA.
ID Q9D6Q6
AC Q9D6Q6
DT 01-JUN-2001 (TREMBlrel. 17, Created)
DT 01-JUN-2001 (TREMBlrel. 17, Last sequence update)
DT 01-MAR-2003 (TREMBlrel. 23, Last annotation update)
DE Adult male tongue cDNA, RIKEN full-length enriched library,
DE clone:2310067K20, full insert sequence.
GN 2310051B21RIK.
OS Mus musculus (Mouse).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
OX NCBI_TaxID=10090;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=C57BL/6J; TISSUE=Tongue;
RC MEDLINE=21085660; PubMed=11217851;
RA Kawai J., Shinagawa A., Shibata K., Yoshino M., Itoh M., Ishii Y.,
RA Arakawa T., Hara A., Fukunishi Y., Konno H., Adachi J., Fukuda S.,
RA Aizawa K., Izawa M., Nishi K., Kiyosawa H., Kondo S., Yamanaka I.,
RA Saïto T., Okazaki Y., Gojōbōri T., Bono H., Kasukawa T., Saïto R.,
RA Kadota K., Matsuda H.A., Ashburner M., Batalov S., Casavant T.,
RA Fleischmann W., Gaasterland T., Giesi C., King B., Kochiwa H.,
RA Kuehl P., Lewis S., Matsuo Y., Nikaïdo I., Pesole G., Quackenbush J.,
RA Schriml L.M., Staudl F., Suzuki R., Tomita M., Wagner L., Washio T.,
RA Blake J., Okido T., Furuno M., Aono H., Baldarelli R., Barsh G.,
RA Brownstein M.J., Bult C., Fletcher C., Fujita M., Gariboldi M.,
RA Gustincich S., Hill D., Hofmann M., Hume D.A., Kamiya M., Lee N.H.,
RA Lyons P., Marchionni L., Mashima J., Mazzarelli J., Mombaerts P.,
RA Nordone P., Ringwald M., Rodriguez I., Sakamoto N.,
RA Sasaki H., Sato K., Schoenbach C., Seya T., Shibata Y., Storch K.-F.,
RA Suzuki H., Toyō-oka K., Wang K.H., Weitz C., Whittaker C., Wilming L.,
RA Wynshaw-Boris A., Yoshida K., Hasegawa Y., Kawaji H., Kohatsu S.,
RA Hayashizaki Y.,
RT "Functional annotation of a full-length mouse cDNA collection.";
RL Nature 409:685-690(2001).
DR EMBL; AK010093; BAB26697.1; -.
DR HSSP; P07098; 1HLG.
DR MGD; MGI:1914967; 2310051B21RIK.
DR InterPro; IPR000073; A/b_hydrolase.
DR InterPro; IPR000734; Lipase.
DR InterPro; IPR000379; Ser_estrs_site.
DR Pfam; PF00561; abhydrolase; 1.
DR PROSITE; PS00120; LIPASE_SER; 1.
SQ SEQUENCE 395 AA; 44623 MW; 245055F5E7FF07C9 CRC64;

Query Match 50.5%; Score 1134.5; DB 11; Length 395;
Best Local Similarity 53.8%; Pred. No. 2.8e-95;
Matches 214; Conservative 69; Mismatches 108; Indels 7; Gaps 3;

QY 17 MWLLILVAMFQ-RNVNSVHMTKAVDPEAFNISEIIHQGYPCPEEVEVATEDGYILSV 75
DB 1 MWLLVTSVLSAFGNGHGLFGKLPKNPEANNVSMITYGYSEEVVTEDEGYILGV 60

QY 76 NRIPGLVQPKTKGRPVVLLQHLGVGGASNWSLNNPNSLGFILADAGFVWNGNSRGN 135
DB 61 YRIPYKKNSENIGRPVAYLQHLGVASATNWTNLPNSLAFILADAGYDVLGNSRGN 120

QY 136 AWSRKHKLISIDQDFWAFSDYEMARFDPVAVINFILOKTGOEKIYVGYSGTMTGFIA 195
DB 121 TWSRKNVYSPDSVFWAFSDFEMAKYDLPATIDFIVQKTGOEKIHYVGHSGTITGIFA 180

QY 196 FSTMPALAKIKRYPALAPATVTKHAKSPGTFKLLPDMMIKGLFGKKEFL-...YQTRFL 252
DB 181 FSTNPALAKIKRYPALAPATVTKYTESPFKKISLIPKFLKLVIFGNKWFPHNYLQDFL 240

QY 253 ROLVYLCGVILDOICSNIMLLGGFTNNWNSRASYAAHLAGTSVQNLHWSQAV 312
DB 241 GTEV---CSRELLDLCSNAULFCGFDKKNLNSRFDVYLGHNPA GSTQDLFHAQLA 297

```

```

QY 196 FSTMPBLAQKIMYPALAPATVYKAKSPGTKEFLLPDMWIKGLFGKKEPL---YOTRFL 252
DB 181 FSTNPALAKKIRFRVALAPATVYKAKSPGTKEFLLPDMWIKGLFGKKEPL---YOTRFL 240
QY 253 ROLVYLCQVILDOICSNIMLLGCFNTNNMNSRASVYAAHTLAGTSVONILHWSQAV 312
DB 241 GTEV---CSRELLDLCSNALFICGDFCKKNLNVSRFDVYLGHNPAQSTQDLFWAQLA 297
QY 313 NSGELRAFQWGSSEKCNKLEKCNQPTPVRYRVRDMTPTAMTGGQWLNPNEDVKNLSEV 372
DB 298 KSGKLAQVNWGSPLRNLHYNQKTPPYDYVSANTVPIAVWNGGHDLADPQDVAMLLPKL 357
QY 373 TNLIVYHKNIPWAHVDVFWGLDAPHEWYNEIHLMOQE 410
DB 358 PNLLYHKEILPYNHLDFIWMADAPQEVYNEIVTMMAED 395

RESULT 8
Q9D6X0 PRELIMINARY; PRT; 395 AA.
AC Q9D6X0;
DT 01-JUN-2001 (TREMELrel. 17, Created)
DT 01-JUN-2001 (TREMELrel. 17, Last sequence update)
DT 01-MAR-2003 (TREMELrel. 23, Last annotation update)
DE 2310051B21RIK protein.
GN 2310051B21RIK.
OS Mus musculus (Mouse).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
OX NCBI_TaxID=10090;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=C57BL/6J; TISSUE=Tongue;
RX MEDLINE=21085660; PubMed=11217851;
RA Kawai J., Shingawa A., Shibata K., Yoshino M., Itoh M., Ishii Y.,
RA Arakawa T., Hara A., Fukunishi Y., Konno H., Adachi J., Fukuda S.,
RA Aizawa K., Izawa M., Nishi K., Kiyosawa H., Kondo S., Yamanaka I.,
RA Saito T., Okazaki Y., Gojobori T., Bono H., Kasukawa T., Saito R.,
RA Kadota K., Matsuda H.A., Ashburner M., Batalov S., Casavant T.,
RA Fleischmann W., Gaasterland T., Glasi C., King B., Kochiwa H.,
RA Kuehl P., Lewis S., Matsuo Y., Nikaide I., Pesole G., Quackenbush J.,
RA Schriml L.M., Staubli F., Suzuki R., Tomita M., Wagner L., Washio T.,
RA Sakai K., Okido T., Furuno M., Aono H., Baldarelli R., Barsh G.,
RA Blake J., Boffelli D., Bojunga N., Carninci P., de Bonaldo M.F.,
RA Brownstein M.J., Bult C., Fletcher C., Fujita M., Gariboldi M.,
RA Gustincich S., Hill D., Hofmann M., Hume D.A., Kamiya M., Lee N.H.,
RA Lyons P., Marchionni L., Makhina J., Mazzarelli J., Mombaerts P.,
RA Nordone P., Ring B., Ringwald M., Rodriguez I., Sakamoto N.,
RA Sasaki H., Sato K., Schoenbach C., Seya T., Shibata Y., Storch K.-F.,
RA Suzuki H., Toyooka K., Wang K.H., Weitz C., Whittaker C., Wilming L.,
RA Wynshaw-Boris A., Yoshida K., Hasegawa Y., Kawai H., Kohtsuki S.,
RA Hayashizaki Y.;
RT "Functional annotation of a full-length mouse cDNA collection.";
RL Nature 409:685-690(2001).
DR EMBL; AK009875; BAB26556.1; -
DR HSSP; P07038; LHG.
DR MCD; MG1:1514967; 2310051B21RIK.
DR InterPro; IPR000073; A/b hydrolase.
DR InterPro; IPR000734; Lipase.
DR InterPro; IPR000379; Ser esterase.
DR Pfam; PF00561; abhydrolase; 1.
DR PROSITE; PS00120; LIPASE SER; 1.
SQ SEQUENCE 395 AA; 44665 MW; 40CNA6B67859A8C5B CRC64;

Query Match 50.5%; Score 1134.5; DB 11; Length 395;
Best Local Similarity 53.5%; Pred. No. 2.8e-95;
Matches 213; Conservative 70; Mismatches 108; Indels 7; Gaps 3;

QY 17 MWLLILVAYMFO-RNVNSVHMTKAVDPEAFNISEITQHQPCEEYEVATEDGYILSV 75
DB 1 MWLLILVTSLSAAGGAGLFGKLGPKNPANNVSMITWYGPSEYEVATEDGYILGV 60

```

```

QY 76 NRIPRGLVOPKKTGSRPVLLQHLGVGASWISLNNPNSLGFILADAGFDVWNGSRGN 135
DB 61 YRIPYKGNSENITGKRPVAYLOHGLIASATNTNLNPNNSLAFILADAGFDVWNGSRGN 120
QY 136 AWSRKHTKLTISIDQDEFWAFSYDEMAFDFLPVAVINFLQKTGQEKIYVGVSGQTTFGFA 195
DB 121 TWSRKNVYSPDSVETWAFSYDEMAFYDLPATDFVQKTGQEKIYVGVSGQTTFGFA 180
QY 196 FSTMPBLAQKIMYPALAPATVYKAKSPGTKEFLLPDMWIKGLFGKKEPL---YOTRFL 252
DB 181 FSTNPALAKKIRFRVALAPATVYKAKSPGTKEFLLPDMWIKGLFGKKEPL---YOTRFL 240
QY 253 ROLVYLCQVILDOICSNIMLLGCFNTNNMNSRASVYAAHTLAGTSVONILHWSQAV 312
DB 241 GTEV---CSRELLDLCSNALFICGDFCKKNLNVSRFDVYLGHNPAQSTQDLFWAQLA 297
QY 313 NSGELRAFQWGSSEKCNKLEKCNQPTPVRYRVRDMTPTAMTGGQWLNPNEDVKNLSEV 372
DB 298 KSGKLAQVNWGSPLRNLHYNQKTPPYDYVSANTVPIAVWNGGHDLADPQDVAMLLPKL 357
QY 373 TNLIVYHKNIPWAHVDVFWGLDAPHEWYNEIHLMOQE 410
DB 358 PNLLYHKEILPYNHLDFIWMADAPQEVYNEIVTMMAED 395

RESULT 9
Q9C9P7 PRELIMINARY; PRT; 395 AA.
AC Q9C9P7;
DT 01-JUN-2001 (TREMELrel. 17, Created)
DT 01-JUN-2001 (TREMELrel. 17, Last sequence update)
DT 01-MAR-2003 (TREMELrel. 23, Last annotation update)
DE 2310051B21RIK protein.
GN 2310051B21RIK.
OS Mus musculus (Mouse).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
OX NCBI_TaxID=10090;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=C57BL/6J; TISSUE=Tongue;
RX MEDLINE=21085660; PubMed=11217851;
RA Kawai J., Shingawa A., Shibata K., Yoshino M., Itoh M., Ishii Y.,
RA Arakawa T., Hara A., Fukunishi Y., Konno H., Adachi J., Fukuda S.,
RA Aizawa K., Izawa M., Nishi K., Kiyosawa H., Kondo S., Yamanaka I.,
RA Saito T., Okazaki Y., Gojobori T., Bono H., Kasukawa T., Saito R.,
RA Kadota K., Matsuda H.A., Ashburner M., Batalov S., Casavant T.,
RA Fleischmann W., Gaasterland T., Glasi C., King B., Kochiwa H.,
RA Kuehl P., Lewis S., Matsuo Y., Nikaide I., Pesole G., Quackenbush J.,
RA Schriml L.M., Staubli F., Suzuki R., Tomita M., Wagner L., Washio T.,
RA Sakai K., Okido T., Furuno M., Aono H., Baldarelli R., Barsh G.,
RA Blake J., Boffelli D., Bojunga N., Carninci P., de Bonaldo M.F.,
RA Brownstein M.J., Bult C., Fletcher C., Fujita M., Gariboldi M.,
RA Gustincich S., Hill D., Hofmann M., Hume D.A., Kamiya M., Lee N.H.,
RA Lyons P., Marchionni L., Makhina J., Mazzarelli J., Mombaerts P.,
RA Nordone P., Ring B., Ringwald M., Rodriguez I., Sakamoto N.,
RA Sasaki H., Sato K., Schoenbach C., Seya T., Shibata Y., Storch K.-F.,
RA Suzuki H., Toyooka K., Wang K.H., Weitz C., Whittaker C., Wilming L.,
RA Wynshaw-Boris A., Yoshida K., Hasegawa Y., Kawai H., Kohtsuki S.,
RA Hayashizaki Y.;
RT "Functional annotation of a full-length mouse cDNA collection.";
RL Nature 409:685-690(2001).
DR EMBL; AK010236; BAB26787.1; -
DR EMBL; AK009413; BAB26272.1; -
DR EMBL; AK009428; BAB26280.1; -
DR EMBL; AK009459; BAB26300.1; -
DR EMBL; AK009473; BAB26312.1; -
DR EMBL; AK009474; BAB26313.1; -
DR EMBL; AK009479; BAB26316.1; -
DR EMBL; AK009523; BAB26338.1; -
DR EMBL; AK009525; BAB26339.1; -
DR EMBL; AK009546; BAB26352.1; -
DR EMBL; AK009571; BAB26368.1; -

```

```

DR EMBL; AK009573; BAB26370.1; -
DR EMBL; AK009723; BAB26466.1; -
DR EMBL; AK009773; BAB26495.1; -
DR EMBL; AK010019; BAB26647.1; -
DR EMBL; AK010035; BAB26656.1; -
DR EMBL; AK010058; BAB26673.1; -
DR EMBL; AK010061; BAB26675.1; -
DR EMBL; AK010124; BAB26715.1; -
DR EMBL; AK010145; BAB26716.1; -
DR HSSP; P07098; 1HLG.
DR MGD; MGI:1914967; 2310051B21Rik.
DR InterPro; IPR000073; A/b hydrolase.
DR InterPro; IPR000734; Lipase.
DR Pfam; PF00561; abhydrolase_1_site.
DR PROSITE; PS00120; LIPASE SER; 1.
SQ SEQUENCE 395 AA; 44637 MW; D3P96B65EA671E34 CRC64;

Query Match 50.4%; Score 1133.5; DB 11; Length 395;
Best Local Similarity 53.5%; Pred. No. 3.4e-95;
Matches 213; Conservative 70; Mismatches 108; Indels 7; Gaps 3;

QY 17 MWLLILVAYMFQ-RNVNSVHPTKAVDPEAPNISEIIHQGYPCSEYEVATEDGYILSV 75
DB 1 MWLLLVTSVLSAFGGAHGLFGKLGPKNPEANMNVSMITYGYPSEYEVATEDGYILV 60

QY 76 NRIPRGLVQPKTGRPVVLLQHLGVGASWISNLPNSLGFILADAGFDVWNGSRGN 135
DB 61 YRIPYKKNSENIQKRPVAYLQHLIASATWITNLPNSLAFILADAGYDVWNGSRGN 120

QY 136 AWSRKHKLTSLDQDEFWAFSDENARFDPVAVNFILQKQGEKIYVYVSGQTTMGFIA 195
DB 121 TWSRKNVYSPDSVEFWAFSDENAKYDLPATIDFIVQKTQGEKIHYVGHSGQTTIGFIA 180

QY 196 FSTWPELAQKIMYFALAPIATVHAKSPGKFKLLPDMIKLFGKKEFL--YQTRFL 252
DB 181 FSTNPAKAKIKRFVALAPVATVYTESPFKKISLIPKLLKLVIFGNMFPNHYLDQFL 240

QY 253 RQLVYLCCQVILDOICSNIMLLGGFNTNNMNSRASVYAAHTLAGTSVQNLHWSQAV 312
DB 241 GTEV---CSRELLDLCSNALFIFCGFDKKNLVSRFGVILGHPAGTSTQDLFWAQLA 297

QY 313 NSGELRAFDMGSETKLEKCNQPTPVRYRVDMTVPTAMTGGODWLSNEDVQMLLSEV 372
DB 298 KSGKLQAVNWGSPQNMLHYNQKTPPYDYDSAMTVPIAVWNGGHDILADPDQVAMLLPKL 357

QY 373 TNLIVHKNIPWAHVDFIWLGLDAPHRMYNEIHLMOQE 410
DB 358 PNLLYHKEILPYNHLDFIWDADAPQEVNEIVTWMAED 395

RESULT 10
Q9C9P8 PRELIMINARY; PRT; 395 AA.
AC Q9C9P8
DT 01-JUN-2001 (TREMELrel. 17, Created)
DT 01-JUN-2001 (TREMELrel. 17, Last sequence update)
DT 01-MAR-2003 (TREMELrel. 23, Last annotation update)
DE Adult male tongue cDNA, RIKEN full-length enriched library,
DE clone:23100761513, full insert sequence.
GN 2310051B21Rik.
OS Mus musculus (Mouse).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
OX NCBI_TaxID=10090;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=C57BL/6J; TISSUE=Tongue;
RX MEDLINE=21085660; PubMed=11217851;
RA Kawai J., Shinagawa A., Shibata K., Yoshino M., Itoh M., Ieshii Y.,
RA Aizawa T., Hara A., Fukunishi Y., Konno H., Adachi J., Fukuda S.,
RA Arizawa K., Izawa M., Nishi K., Kiyosawa H., Kondo S., Yamanaka I.,
RA Saico T., Okazaki Y., Gojibori T., Bono H., Kasukawa T., Saito R.,
RA Kadota K., Matsuda H.A., Ashburner M., Batalov S., Casavant T.,

```

```

RA Fleischmann W., Caasterland T., Gissi C., King B., Kochiwa H.,
RA Kuehl P., Lewis S., Mateo Y., Nikaio I., Pesole G., Quackenbush J.,
RA Schriml L.M., Staubli F., Suzuki R., Tomita M., Wagner L., Washio T.,
RA Sakai K., Okido T., Furuno M., Aono H., Baldarelli R., Barh G.,
RA Blake J., Boffelli D., Bojunga N., Carninci P., de Bonaldo M.F.,
RA Brownstein M.J., Bult C., Fletcher C., Fujita M., Gariboldi M.,
RA Gustincich S., Hill D., Hofmann M., Hume D.A., Kamiya M., Lee N.H.,
RA Lyons P., Marchionni L., Mashima J., Mazzarelli J., Nombaerts P.,
RA Nordone P., Ring B., Ringwald M., Rodriguez I., Sakamoto N.,
RA Suzuki H., Sato K., Schoenbach C., Seya T., Shibata Y., Storch K.-F.,
RA Wynshaw-Boris A., Yoshida K., Hasegawa Y., Kawaji H., Kohetsuki S.,
RA Hayashizaki Y.;
RT "functional annotation of a full-length mouse cDNA collection.";
RL Nature 409:685-690(2001).
DR EMBL; AK010116; BAB26711.1; -
DR EMBL; AK009300; BAB26201.1; -
DR HSSP; P07098; 1HLG.
DR MGD; MGI:1914967; 2310051B21Rik.
DR InterPro; IPR000073; A/b hydrolase.
DR InterPro; IPR000734; Lipase.
DR InterPro; IPR000379; Ser_estra_site.
DR Pfam; PF00561; abhydrolase_1.
DR PROSITE; PS00120; LIPASE SER; 1.
SQ SEQUENCE 395 AA; 44579 MW; D2296865EA671D34 CRC64;

Query Match 50.4%; Score 1133.5; DB 11; Length 395;
Best Local Similarity 53.5%; Pred. No. 3.4e-95;
Matches 213; Conservative 70; Mismatches 108; Indels 7; Gaps 3;

QY 17 MWLLILVAYMFQ-RNVNSVHPTKAVDPEAPNISEIIHQGYPCSEYEVATEDGYILSV 75
DB 1 MWLLLVTSVLSAFGGAHGLFGKLGPKNPEANMNVSMITYGYPSEYEVATEDGYILV 60

QY 76 NRIPRGLVQPKTGRPVVLLQHLGVGASWISNLPNSLGFILADAGFDVWNGSRGN 135
DB 61 YRIPYKKNSENIQKRPVAYLQHLIASATWITNLPNSLAFILADAGYDVWNGSRGN 120

QY 136 AWSRKHKLTSLDQDEFWAFSDENARFDPVAVNFILQKQGEKIYVYVSGQTTMGFIA 195
DB 121 TWSRKNVYSPDSVEFWAFSDENAKYDLPATIDFIVQKTQGEKIHYVGHSGQTTIGFIA 180

QY 196 FSTWPELAQKIMYFALAPIATVHAKSPGKFKLLPDMIKLFGKKEFL--YQTRFL 252
DB 181 FSTNPAKAKIKRFVALAPVATVYTESPFKKISLIPKLLKLVIFGNMFPNHYLDQFL 240

QY 253 RQLVYLCCQVILDOICSNIMLLGGFNTNNMNSRASVYAAHTLAGTSVQNLHWSQAV 312
DB 241 GTEV---CSRELLDLCSNALFIFCGFDKKNLVSRFGVILGHPAGTSTQDLFWAQLA 297

QY 313 NSGELRAFDMGSETKLEKCNQPTPVRYRVDMTVPTAMTGGODWLSNEDVQMLLSEV 372
DB 298 KSGKLQAVNWGSPQNMLHYNQKTPPYDYDSAMTVPIAVWNGGHDILADPDQVAMLLPKL 357

QY 373 TNLIVHKNIPWAHVDFIWLGLDAPHRMYNEIHLMOQE 410
DB 358 PNLLYHKEILPYNHLDFIWDADAPQEVNEIVTWMAED 395

RESULT 11
Q9D6L9 PRELIMINARY; PRT; 395 AA.
AC Q9D6L9
DT 01-JUN-2001 (TREMELrel. 17, Created)
DT 01-JUN-2001 (TREMELrel. 17, Last sequence update)
DT 01-MAR-2003 (TREMELrel. 23, Last annotation update)
DE Adult male tongue cDNA, RIKEN full-length enriched library,
DE clone:23100761513, full insert sequence.
GN 2310051B21Rik.
OS Mus musculus (Mouse).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
OX NCBI_TaxID=10090;

```

RESULT 12	
Q9D6T5	
ID Q9D6T5	PRELIMINARY;
AC Q9D6T5	PRT; 395 AA.
DT 01-JUN-2001	(TRENBLrel. 17, Created)
DT 01-JUN-2001	(TRENBLrel.17, Last sequence update)


```

313 NSGELRAFDWGSEKXLEKCKNQPTVRYRVEDMTVTPTAMWTGGODLNSNPEDVKKLLSEV 372
398 KSGKLOAYNWGSPFQNMHLHYKTPPYDYDSAMTVPFAVWNGGHDILADPDQVAKLLPKL 357
373 TNLIIYHKNIPWAHVDFIWLGDAPHRMYNEIHLMOQE 410
358 PNLLYHKEILPYNHLDFIWMADPOEVNEIWTWMAED 395

RESULT 14
Q9D7C5 PRELIMINARY; PRT; 395 AA.
ID Q9D7C5;
AC Q9D7C5;
DT 01-JUN-2001 (TRENBLrel. 17, Created)
DT 01-JUN-2001 (TRENBLrel. 17, Last sequence update)
DT 01-MAR-2003 (TRENBLrel. 23, Last annotation update)
DE 2310051B21RIK protein.
DE 2310051B21RIK.
OS Mus musculus (Mouse).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
OX NCBI_TaxID=10090;
RN [1]
SEQUENCE FROM N.A.
RC STRAIN=C57BL/6J; TISSUE=Tongue;
RX MEDLINE=21085660; PubMed=11217851;
RA Kawai J., Shinagawa A., Shibata K., Yoshino M., Itoh M., Ishii Y.,
RA Arakawa T., Hara A., Fukunishi Y., Konno H., Adachi J., Fukuda S.,
RA Aizawa K., Izawa M., Nishi K., Kiyosawa H., Kondo S., Yamanaka I.,
RA Saito T., Okazaki Y., Gojohori T., Bono H., Kasukawa T., Saito R.,
RA Kadota K., Matsuda H.A., Ashburner M., Batalov S., Casavant T.,
RA Fleischmann W., Gaasterland T., Gissi C., King B., Kochiwa H.,
RA Kuehl P., Lewis S., Matsuo Y., Nikaide I., Pesole G., Quackenbush J.,
RA Schriml L.M., Staubli F., Suzuki R., Tomita M., Wagner L., Washio T.,
RA Sakai K., Okido T., Furuno M., Aono H., Baldarelli R., Barsh G.,
RA Blake J., Boffelli D., Bojunga N., Carninci P., de Bonaldo M.F.,
RA Brownstein M.J., Bult C., Fletcher C., Fujita M., Gariboldi M.,
RA Gustincich S., Hill D., Hofmann M., Hume D.A., Kamiya M., Lee N.H.,
RA Lyons P., Marchionni L., Mashima J., Mazzarelli J., Mombaerts P.,
RA Nordone P., Ring B., Ringwald M., Rodriguez I., Sakamoto N.,
RA Sasaki H., Sato K., Schoenbach C., Seya T., Shibata Y., Storch K.-F.,
RA Suzuki H., Toyooka K., Wang K.H., Weitz C., Whittaker C., Wilming L.,
RA Wynshaw-Boris A., Yoshida K., Hasegawa Y., Kawaji H., Kohtsuki S.,
RA Hayashizaki Y.,
RA "Functional annotation of a full-length mouse cDNA collection."
RT Nature 409:685-690(2001).
DR EMBL; AK009537; BAB26346.1; -.
DR HSSP; P07098; 1HLG.
DR MGD; MGI:1914967; 2310051B21RIK.
DR InterPro; IPR000073; A/b hydrolase.
DR InterPro; IPR000734; Lipase.
DR Pfam; PF00561; abhydrolase_1.
DR PROSITE; PS00120; LIPASE SR; 1.
SQ SEQUENCE 395 AA; 44607 MW; E43317C2254FA8FB CRC64;

Query Match 50.2%; Score 1128.5; DB 11; Length 395;
Best Local Similarity 53.3%; Pred. No. 9.8e-95;
Matches 212; Conservative 70; Mismatches 109; Indels 7; Gaps 3;

QY 17 MWLLILVAYMFQ-RNVNSVHMFPTKAVDPEAFMNISEIIHQGYPCPEEYEVATEDGYILSV 75
DB 1 MWLLVTSVLSAFGAGHLFGKLGPKNPEANMNSQMITYGYPSEEVVTEGYILGV 60

QY 76 NRIPGLVQPKTSRPPVLLQHLGVGGASNIWLNPNLSLGLFADAGFDVWNGSRGN 135
DB 61 YRIPYKKNSENIGKRPVAYLQHLGASATNITLNPNSLAFIADAGYDVLNLSRGN 120

QY 136 AMSRKHKTLISIDQDEFWAFSPDEMARFOLPAVINFILOKTQEKIYVYVSGYTMGFGIA 195
DB 121 TWSRKNVYSPDSVEFWAFSPDEMAKYDLPATIDFIVQKTQEKIHYVGHSGTIGFIA 180

QY 196 FSTMPELLAQIKMYFALAPIATVTKAKSPGTFKLLPDMIMKGLFGKKEFL---YQTRFL 252
DB 181 FSTNPALAKIKRFRFALAPVATVYTESPFKKISLIPKFLKLVIFGNKCFMHPHYLDQFL 240

QY 253 RQLVLYLQGVITDQICNIMLLGLGFNTNNMNSRASAAYAAHTLAGTSVQNLHWSQAV 312
DB 241 GTEV---CSRLLDILCSNALFIFCGFDKKNLNSRFDVYLGHNPAGTSTQDLFFHQA 297

```



```

Db 181 FSTPALAKKIRFYALAPVATVYTESPFKKHIFPKFLKVI FGNKMFPHNYLDQFL 240
QY 253 RQLVYLCOGQVILDOICSNIMLLGGFNTNNMNSRASVYAAHTLAGTSVQNILHWSQAV 312
Db 241 GTEV---CSRELLDLCSNLFIFCGPDKKLVSRPVDVILGHNPAGTSTQDLFPHWAQLA 297
QY 313 NSGELRAPDWGSETNLEKCNQPTPVRYRVRDMTPTAMWTGGQDWLSNPDVVKLLSEV 372
Db 298 KSGKLQAYNWGSPQLNMLHYNQKTPPYDYVSAMTPTAVWNGGHDILADPDQVAMLLPKL 357
QY 373 TNLVYHKNIPEWAHVDFTWGLDAPHRMNEIHLMOQE 410
Db 358 PNLLYHKEILPYNHLDFIWMADAPQEVYNEIVTMAED 395

RESULT 15
Q9D6L1
ID Q9D6L1 PRELIMINARY; PRT; 395 AA.
AC Q9D6L1
DC Q9D6L1
DT 01-JUN-2001 (TRENDELrel. 17, Created)
DT 01-JUN-2001 (TRENDELrel. 17, Last sequence update)
DT 01-MAR-2003 (TRENDELrel. 23, Last annotation update)
DE Adult male tongue cDNA, RIKEN full-length enriched library,
DE clone:2310079Q20, full insert sequence.
GN 2310051B21RIK.
OS Mus musculus (Mouse).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
OX NCBI_TaxID=10090;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=C57BL/6J; TISSUE=Tongue;
RX MEDLINE=21085660; PubMed=11217851;
RA Kawai J., Shinagawa A., Shibata K., Yoshino M., Itoh M., Ishii Y.,
RA Arakawa T., Hara A., Fukunishi Y., Konno H., Adachi J., Fukuda S.,
RA Aizawa K., Izawa M., Nishi K., Kiyosawa H., Kondo S., Yamanaka I.,
RA Saito T., Okazaki Y., Gojobori T., Bono H., Kasukawa T., Saito R.,
RA Kadota K., Matsuda H.A., Ashburner M., Batalov S., Casavant T.,
RA Pleischmann W., Gaasterland T., Giesi C., King B., Kochiwa H.,
RA Kuehl P., Lewis S., Matsuo Y., Nikaido I., Pesole G., Quackenbush J.,
RA Schriml L.M., Staubli P., Suzuki R., Tomita M., Wagner L., Washio T.,
RA Sakai K., Okido T., Furuno M., Aono H., Baldarelli R., Barsh G.,
RA Blake J., Boffelli D., Bojunga N., Carninci P., de Bonaldo M.F.,
RA Brownstein M.J., Bult C., Fletcher C., Fujita M., Gariboldi M.,
RA Gustincich S., Hill D., Hofmann M., Hume D.A., Kamiya M., Lee N.H.,
RA Lyons P., Marchionni L., Mashima J., Mazzarelli J., Mombaerts P.,
RA Nordone P., Ring B., Ringwald M., Rodriguez I., Sakamoto N.,
RA Sasaki H., Sato K., Schoenbach C., Seya T., Shibata Y., Storch K.-P.,
RA Suzuki H., Toyooka K., Wang K.H., Weitz C., Whitaker C., Wilming L.,
RA Wyshaw-Boris A., Yoshida K., Hasegawa Y., Kawaji H., Kohtsuki S.,
RA Hayashizaki Y.;
RT "Functional annotation of a full-length mouse cDNA collection.";
RL Nature 409:685-690(2001).
DR EMBL; AK010231; BAB26784.1; -.
DR HSSP; P07098; iHLG.
DR MGD; MGI:1914967; 2310051B21RIK.
DR InterPro; IPR000073; A/b hydrolase.
DR InterPro; IPR000734; Lipase.
DR InterPro; IPR000379; Ser esters site.
DR Pfam; PF00561; abhydrolase; 1_-site.
DR PROSITE; PS00120; LIPASE SER; 1.
SQ SEQUENCE 395 AA; 44659 MW; 13E0BB95FC5A40E0 CRC64;

```

```

Query Match 50.1%; Score 1126.5; DB 11; Length 395;
Best Local Similarity 53.5%; Pred. No. 1.5e-94;
Matches 213; Conservative 69; Mismatches 109; Indels 7; Gaps 3;

QY 17 MWLLILVAVNQ-RNVNSVHMFKAVDPEAFNISEIICHQYCPCEEEYEVATEDGYILSV 75
Db 1 MWLLLVTSLSAFGGAGHGFGLGPKNPENANNVSMQITWYGPSEYEVVTEGDIILGV 60
QY 76 NRIPRGLVQPKTGTSPVLLQLHGLVGGASNWSNLNPNNSLGFILADAGFDVWNGNSRGN 135

```

```

Db 61 YRIPYKKNSENIKRPPVAYLOHGLIASATNITNLPNNSLAFILADAGYDVLGNSRGN 120
QY 136 AWSRKHKTLSDIDQDEFWAFSPYDEMARFDLPVAVINFILQKTOEKIYVYVSGQTTMGFIA 195
Db 121 TWSRKNVYSPDSVEFWAFSPDEMAKYDLPATIDFIVQKTQOEKIHVYVHSGQTTIGFIA 180
QY 196 FSTMPDLAQKIKMYPALAPIATVKAHSPGTFKLLPDMMIKGLFGKKEFL---YQTRPL 252
Db 181 FSTPALAKKIRFYALAPVATVYTESPFKKHIFPKFLKVI FGNKMFPHNYLDQFL 240
QY 253 RQLVYLCOGQVILDOICSNIMLLGGFNTNNMNSRASVYAAHTLAGTSVQNILHWSQAV 312
Db 241 GTEV---CSRELLDLCSNLFIFCGPDKKLVSRPVDVILGHNPAGTSTQDLFPHWAQLA 297
QY 313 NSGELRAPDWGSETNLEKCNQPTPVRYRVRDMTPTAMWTGGQDWLSNPDVVKLLSEV 372
Db 298 KSGKLQAYNWGSPQLNMLHYNQKTPPYDYVSAMTPTAVWNGGHDILADPDQVAMLLPKL 357
QY 373 TNLVYHKNIPEWAHVDFTWGLDAPHRMNEIHLMOQE 410
Db 358 PNLLYHKEILPYNHLDFIWMADAPQEVYNEIVTMAED 395

```

Search completed: February 19, 2004, 14:22:24
Job time : 35.1771 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2004 Compugen Ltd.

OM nucleic - nucleic search, using sw model

Run on: February 19, 2004, 04:41:57 ; Search time 7514.62 Seconds
(without alignments)
11127.553 Million cell updates/sec

Title: US-10-042-431-45
Perfect score: 2044
Sequence: 1 gtcgaccacgcgtccgggg.....aaaaaaagggcgccgc 2044

Scoring table: IDENTITY_NUC
Gapop 10.0 , Gapext 1.0

Searched: 2888711 seqs, 2045481386 residues

Total number of hits satisfying chosen parameters: 5777422

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database :

GenEmbl.*

- 1: gb_ba.*
- 2: gb_htg.*
- 3: gb_in.*
- 4: gb_on.*
- 5: gb_ov.*
- 6: gb_pat.*
- 7: gb_ph.*
- 8: gb_pl.*
- 9: gb_pr.*
- 10: gb_ro.*
- 11: gb_sts.*
- 12: gb_sy.*
- 13: gb_un.*
- 14: gb_vl.*
- 15: em_ba.*
- 16: em_fun.*
- 17: em_hum.*
- 18: em_in.*
- 19: em_mu.*
- 20: em_on.*
- 21: em_or.*
- 22: em_ov.*
- 23: em_pat.*
- 24: em_ph.*
- 25: em_pl.*
- 26: em_ro.*
- 27: em_sts.*
- 28: em_un.*
- 29: em_vl.*
- 30: em_htg_hum.*
- 31: em_htg_inv.*
- 32: em_htg_other.*
- 33: em_htg_mus.*
- 34: em_htg_pin.*
- 35: em_htg_rod.*
- 36: em_htg_mam.*
- 37: em_htg_vrt.*
- 38: em_sy.*
- 39: em_htgo_hum.*
- 40: em_htgo_mus.*
- 41: em_htgo_other.*

Pred. No. is the number of results predicted by chance to have a

score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Match	Length	ID	Description
1	1138	55.7	1210	6	AX521798 Sequence
2	1107.8	54.2	1272	6	AX477670 Sequence
3	1000.4	48.9	2577	10	BC031933 Mus muscu
					BC031933 Human DNA
4	892	43.6	43456	9	AL353113 Homo sapi
5	890.4	43.6	158195	2	AC015461 Homo sapi
6	752	36.8	1116	6	AX521801 Sequence
7	630.2	30.8	76553	2	AC015506 Homo sapi
8	528.2	25.8	2982	9	AX091558 Homo sapi
9	527.8	25.8	2481	6	AX411234 Sequence
10	527.8	25.8	2481	9	HSLYACLY
11	527.8	25.8	2481	9	HSU08464
12	527.8	25.8	2626	9	HSLAL
13	526.2	25.7	2493	6	BD094072 Shear str
14	526.2	25.7	2493	9	HUMLIPCHL
15	524.6	25.7	2586	9	BC012287
16	489.2	23.9	1378	6	A26689
17	489.2	23.9	1378	6	A26690
18	470	23.0	1365	9	HSGLR
19	470	23.0	1367	6	A01046
20	470	23.0	1367	6	A12714
21	467.6	22.9	1336	6	A01157
22	467.6	22.9	1355	10	RNLIP
23	462.2	22.6	2506	9	AK096406
24	461	22.6	2358	10	MMLYACLY
25	459.8	22.5	3144	10	S81497
26	458	22.4	1417	4	BOVME
27	454	22.2	1228	6	AX392821
28	448.4	21.9	1473	6	AX392824
29	446.8	21.9	1185	6	AX472752
30	444.2	21.7	1137	6	AR039022
31	444.2	21.7	1137	6	AR092633
32	444.2	21.7	1140	6	AR039023
33	444.2	21.7	1140	6	AR092634
34	444.2	21.7	1146	6	AR039024
35	444.2	21.7	1146	6	AR092635
36	444.2	21.7	1198	6	A57760
37	444.2	21.7	1528	6	A57756
38	444.2	21.7	1528	6	AR039021
39	444.2	21.7	1528	6	AR092632
40	444.2	21.7	1528	6	AR092632
41	441.8	21.6	1800	5	BC052131
42	441	21.6	1137	6	A39301
43	441	21.6	1140	6	A39303
44	441	21.6	1146	6	A39305
45	441	21.6	1528	6	A39300

ALIGNMENTS

RESULT 1
AX521798
LOCUS AX521798 1210 bp DNA linear PAT 05-OCT-2002
DEFINITION Sequence 1 from Patent WO0236731.
ACCESSION AX521798
VERSION AX521798.1 GI:23572854
KEYWORDS
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.

REFERENCE
1
AUTHORS Xiao Y.
TITLE Human lysosomal acid lipase
JOURNAL Patent: WO 0236731-A 1 10-MAY-2002.

Bayer Aktiengesellschaft (DE)		Location/Qualifiers	
1. .1210		/organism="Homo sapiens"	
/mol_type="genomic DNA"		/db_xref="taxon:9606"	
1. .1197		/note="unnamed protein product"	
/codon_start=1		/protein_id="CAD52970.1"	
/db_xref="GI:23572855"		/translation="MELKLLPRDLVLRWAEITRPLDRKQSEIIHQHQPCEBEYEVATED	
GVLVSNRIPGLVQPKTSRPNVLQHLGVGASWIKVNLNPNLSLGFILLADAGPDV		GVLVSNRIPGLVQPKTSRPNVLQHLGVGASWIKVNLNPNLSLGFILLADAGPDV	
WNGSRGNWRSRKHETLSIDODEFWASDEWAFDLPAVINFILOKTOGEKIYVGY		WNGSRGNWRSRKHETLSIDODEFWASDEWAFDLPAVINFILOKTOGEKIYVGY	
SQGTMTGFIASTHPELAKIKYFALAPIATVYHAKSPGTXKFLLPDMIKGLPKKK		SQGTMTGFIASTHPELAKIKYFALAPIATVYHAKSPGTXKFLLPDMIKGLPKKK	
EFLYTRFLRQLVILVLCQGVILDDQICSNIMLLGGFNTNNMNSRASVYAAHTLUGRS		EFLYTRFLRQLVILVLCQGVILDDQICSNIMLLGGFNTNNMNSRASVYAAHTLUGRS	
VQNIHLWQAVNSGELRAFDMGSETKNLEKNOPTPRYRVRDMTPTVPMWTTGGODWL		VQNIHLWQAVNSGELRAFDMGSETKNLEKNOPTPRYRVRDMTPTVPMWTTGGODWL	
SNPDEVKMLLSEVTLIYHKNIPWAHVDFFIWLDPAPHRMYNEIHLMQOEBETNLSQG		SNPDEVKMLLSEVTLIYHKNIPWAHVDFFIWLDPAPHRMYNEIHLMQOEBETNLSQG	
RCEAVL"			
BASE COUNT	338 a	264 c	296 g 312 t
ORIGIN			
Query Match 55.7%; Score 1138; DB 6; Length 1210;			
Best Local Similarity 100.0%; Pred. No. 5.5e-281;			
Matches 1138; Conservative 0; Mismatches 0; Indels 0; Gaps 0;			
Qy	273	AGTGAATCATCAACATCAAGGCTATCCCTGTGAGGAATATGAAGTCGCAACTGAAGAT	332
Db	73	AGTGAATCATCAACATCAAGGCTATCCCTGTGAGGAATATGAAGTCGCAACTGAAGAT	132
Qy	333	GGGTATATCTTCTGTGTAACAGGATTCCTCGAGGCTAGTGCACACCTAAGAAGACAGGT	392
Db	133	GGGTATATCTTCTGTGTAACAGGATTCCTCGAGGCTAGTGCACACCTAAGAAGACAGGT	192
Qy	393	TCGAGGCTGTGGTGTACTGAGCATGCTAGTGGAGGCTAGTGCACACTGGATTC	452
Db	193	TCGAGGCTGTGGTGTACTGAGCATGCTAGTGGAGGCTAGTGCACACTGGATTC	252
Qy	453	AACTCTCCCAACAATAGCTGGCTTCATTCTGGCAGATGCTGGTTTGGAGTGGATG	512
Db	253	AACTCTCCCAACAATAGCTGGCTTCATTCTGGCAGATGCTGGTTTGGAGTGGATG	312
Qy	513	GGGAACAGCAGGGGAACGCTGGTCTGAAAACACAGACACTCTCCATAGACCAAGAT	572
Db	313	GGGAACAGCAGGGGAACGCTGGTCTGAAAACACAGACACTCTCCATAGACCAAGAT	372
Qy	573	GAGTTCCTGGCTTTCAGTTATGATGAGTGGCTAGTGTTCAGCTTCTCGAGTGAATAAC	632
Db	373	GAGTTCCTGGCTTTCAGTTATGATGAGTGGCTAGTGTTCAGCTTCTCGAGTGAATAAC	432
Qy	633	TTTATTTTGCAGAAAACGGGCCAGGAAAGATCTATTATGTGCGGTATTCACAGGGCACC	692
Db	433	TTTATTTTGCAGAAAACGGGCCAGGAAAGATCTATTATGTGCGGTATTCACAGGGCACC	492
Qy	693	ACCATGGCTTTATTCATTTTCCACCATGCCAGAGCTGGCTCAGAAAATCAAAATGTAT	752
Db	493	ACCATGGCTTTATTCATTTTCCACCATGCCAGAGCTGGCTCAGAAAATCAAAATGTAT	552
Qy	753	TTTGCTTTAGCACCCATAGCCATGTTTAAGCATGCAAAAAGCCCGGACCAAAATTTTG	812
Db	553	TTTGCTTTAGCACCCATAGCCATGTTTAAGCATGCAAAAAGCCCGGACCAAAATTTTG	612
Qy	813	TTGCTGCCAGATATGATCAAGGATTTGTTGGCAAAAAGAAATTTCTGTATCAGACC	872
Db	613	TTGCTGCCAGATATGATCAAGGATTTGTTGGCAAAAAGAAATTTCTGTATCAGACC	672
Qy	873	AGATTTCTCAGACCAACTTTGTTATTTACCTTTTGGCCAGGTGATTTCTGATCAGATTTGT	932
Db	673	AGATTTCTCAGACCAACTTTGTTATTTACCTTTTGGCCAGGTGATTTCTGATCAGATTTGT	732
Qy	933	AGTAATATCATGTTACTTCTGGGTGGATTCACACCAAGATATGACATGACCCGAGCA	992
Db	733	AGTAATATCATGTTACTTCTGGGTGGATTCACACCAAGATATGACATGACCCGAGCA	792

Qy	993	AGTGTATATGCTGCCACACATCTTGTCTGGGAACATCTGTGCAAAATATTCTACACTGGAGC	1053
Db	793	AGTGTATATGCTGCCACACATCTTGTCTGGGAACATCTGTGCAAAATATTCTACACTGGAGC	852
Qy	1053	CAGGCAGTGAATTTCTGGTGAACCTCCGGGCAATTTTGACTGGGGGAGTGAGACCAAAATCTG	1112
Db	853	CAGGCAGTGAATTTCTGGTGAACCTCCGGGCAATTTTGACTGGGGGAGTGAGACCAAAATCTG	912
Qy	1113	GAATAATGCAATCAGCCCAATCTCTGTAAGGTACAGAGTACAGATATGACGGTCCCTACA	1172
Db	913	GAATAATGCAATCAGCCCAATCTCTGTAAGGTACAGAGTACAGATATGACGGTCCCTACA	972
Qy	1173	GCAATGTGCACAGCAGGTTCAGGACTGGCTTTCAAATCCAGAAGACGCTGAAAATGCTGCTC	1232
Db	973	GCAATGTGCACAGCAGGTTCAGGACTGGCTTTCAAATCCAGAAGACGCTGAAAATGCTGCTC	1032
Qy	1233	TCTGAGGTGACCAACCTCATCTACATAAGAAATATTCCTGAATGGGCTCAGCTGGATTTTC	1292
Db	1033	TCTGAGGTGACCAACCTCATCTACATAAGAAATATTCCTGAATGGGCTCAGCTGGATTTTC	1092
Qy	1293	ATCTGGGGTTTGGATGCTCTCCACCGTATGTACAAATGAAATCATCCATCTGATGCAGCAG	1352
Db	1093	ATCTGGGGTTTGGATGCTCTCCACCGTATGTACAAATGAAATCATCCATCTGATGCAGCAG	1152
Qy	1353	GAGGAGACCAACCTTTTCCAGGGACGGTGTGAGCCGCTATTTGTGAAGCATCTGACACT	1410
Db	1153	GAGGAGACCAACCTTTTCCAGGGACGGTGTGAGCCGCTATTTGTGAAGCATCTGACACT	1210

RESULT 2	AX477670	1272 bp	DNA	linear	PAT 12-AUG-2002
LOCUS	Sequence 8 from Patent WO0246418.				
DEFINITION	AX477670				
ACCESSION	AX477670.1				
VERSION	GI:22216841				
KEYWORDS	Homo sapiens (human)				
SOURCE	Homo sapiens				
ORGANISM	Homo sapiens				
REFERENCE	1 Griffin, J.A., Gandhi, A.R., Ramkumar, J., Tang, Y.T., Ding, L., Yue, H., Gietzen, K.J., Sapperstein, S.K., Honchell, C.D., Bruns, C.M., Duggan, B.M., Xu, Y. and Lee, S.				
AUTHORS	Lipid-associated molecules				
TITLE	Patent: WO 0246418-A 8 13-JUN-2002;				
JOURNAL	Incyte Genomics, Inc. (US)				
FEATURES	Location/Qualifiers				
source	1..1272				
	/organism="Homo sapiens"				
	/mol_type="genomic DNA"				
	/db_xref="taxon:9606"				
	/note="Incyte ID No: 7483978CB1"				
BASE COUNT	351 a	263 c	311 g	347 t	
ORIGIN					
Query Match 54.2%; Score 1107.8; DB 6; Length 1272;					
Best Local Similarity 97.2%; Pred. No. 3.1e-273; Indels 21; Gaps 1;					
Matches 1146; Conservative 0; Mismatches 12;					
Qy	240	AAAGCTGTGGACCCAGAACGATTCATGAATATTAGTGAATATCATCAACATCAAGGCTAT	299		
Db	94	AATGAATGATCTCTGAGGTCTGGATGAATAGTGAATCATCAACATCAAGGCTAT	153		
Qy	300	CCCTGTGAGGAATATGAAGTGCACACTGAAGATGGGTATATCTTCTGTGTAAACAGATT	359		
Db	154	CCCTGTGAGGAATATGAAGTGCACACTGAAGATGGGTATATCTTCTGTGTAAACAGATT	213		
Qy	360	CCTCGAGGCTTAGTGCAACCTTAAGAAGACAGGTTCCAGGCTGTGGTGTACTGACGAT	419		
Db	214	CCTCGAGGCTTAGTGCAACCTTAAGAAGACAGGTTCCAGGCTGTGGTGTACTGACGAT	273		

QY 420 GGCTAGTTCGAGTGTCTAGCACTGGATTTCACCTGCTGCCACCAATAGCCTGGCTTC 479
 Db 274 GGCTAGTTCGAGTGTCTAGCACTGGATTTCACCTGCTGCCACCAATAGCCTGGCTTC 333
 QY 480 ATTCTGCAGATGCTGTTTTCAGCTGTGGATGGGAACAGCAGGGGAAACGCTGCTCT 539
 Db 334 ATTCTGCAGATGCTGTTTTCAGCTGTGGATGGGAACAGCAGGGGAAACGCTGCTCT 393
 QY 540 CGAAACACAGACACCTCTCCATAGACCAAGATGAGTTCCTGGCTTTCAGTTATGATGAG 599
 Db 394 CGAAACACAGACACCTCTCCATAGACCAAGATGAGTTCCTGGCTTTCAGTTATGATGAG 453
 QY 600 ATGCTAGTTCAGCTTCCTGCTGATGATAAATTTTATTTGAGAGAAACGGCCAGGAA 659
 Db 454 ATGCTAGTTCAGCTTCCTGCTGATGATAAATTTTATTTGAGAGAAACGGCCAGGAA 513
 QY 660 AAGATCTATTATGTCGGCTATTTCACGGGACCAACCATGGCTTTATTGCAATTTCCACC 719
 Db 514 AAGATCTATTATGTCGGCTATTTCACGGGACCAACCATGGCTTTATTGCAATTTCCACC 573
 QY 720 ATGCCAGAGCTGGCTCAGAAATCAAAATGATTTTCTTTAGCACCATAGCCTGTT 779
 Db 574 ATGCCAGAGCTGGCTCAGAAATCAAAATGATTTTCTTTAGCACCATAGCCTGTT 633
 QY 780 AAGCATGCAAAAGCCCGGGGACCAAAATTTTGTGTCGCCAGATGATGATCAAGGGA 839
 Db 634 AAGCATGCAAAAGCCCGGGGACCAAAATTTTGTGTCGCCAGATGATGATCAAGGGA 693
 QY 840 TTGTTTGGCAAAAGAAATTTCTATCAGACCAAGATTTCTCAGACAACTTGTATTATC 899
 Db 694 TTGTTTGGCAAAAGAAATTTCTATCAGACCAAGATTTCTCAGACAACTTGTATTATC 753
 QY 900 CTTTGTGCCAGGTGATTTCTGATCAGATTTGTAGTAATATCATGTTTACTTCTGGTGA 959
 Db 754 CTTTGTGCCAGGTGATTTCTGATCAGATTTGTAGTAATATCATGTTTACTTCTGGTGA 813
 QY 960 TTCACACCAACATATGACAT-----GAGCCGACGCAAGTGTA 998
 Db 814 TTCACACCAACATATGACATGAACTCATGTTGTTTACGAGCCGACGCAAGTGTA 873
 QY 999 TATGTCGCCACACTTTGCTGGAACATCTGTGCAAAATATTTACACTGAGCCAGGCA 1058
 Db 874 TATGTCGCCACACTTTGCTGGAACATCTGTGCAAAATATTTACACTGAGCCAGGCA 933
 QY 1059 GTGAATTCGTGTAACCTCCGGGATTTTACTGCGGGGAGTGAGACCAAAATCTGAAAAA 1118
 Db 934 GTGAATTCGTGTAACCTCCGGGATTTTACTGCGGGGAGTGAGACCAAAATCTGAAAAA 993
 QY 1119 TGCAATCAGCCAACTCTCTGTAAAGGTACAGAGTACAGATATGACGGTCCCTACAGCAATG 1178
 Db 994 TGCAATCAGCCAACTCTCTGTAAAGGTACAGAGTACAGATATGACGGTCCCTACAGCAATG 1053
 QY 1179 TGAACAGAGTACAGAGTCTGCTTTCAATTCAGACAGCTGAAATGCTGCTCTGAG 1238
 Db 1054 TGAACAGAGTACAGAGTCTGCTTTCAATTCAGACAGCTGAAATGCTGCTCTGAG 1113
 QY 1239 GTGACCAACCTCATCTACCAAGAAATTTCTCTCAATGGGCTCAGCGGATTTTCATCTGG 1298
 Db 1114 GTGACCAACCTCATCTACCAAGAAATTTCTCTCAATGGGCTCAGCGGATTTTCATCTGG 1173
 QY 1299 GGTTCGATGTCCTCACCCTGATGATCAATGAAATCATTCATCTGATGACGAGGAG 1358
 Db 1174 GGTTCGATGTCCTCACCCTGATGATCAATGAAATCATTCATCTGATGACGAGGAG 1233
 QY 1359 ACCAACCTTTCCAGGAGCGGTGTGAGCCGCTATTGTGA 1397
 Db 1234 ACCAACCTTTCCAGGAGCGGTGTGAGCCGCTATTGTGA 1272

RESULT 3
 BC031933
 LOCUS
 DEFINITION Mus musculus, RIKEN cdna 2577 bp mRNA linear ROD 20-SEP-2002
 2577 bp cdna 4632427C23 gene, clone MGC:25790

ACCESSION VERSION KEYWORDS SOURCE ORGANISM

REFERENCE AUTHORS TITLE JOURNAL

REMARK COMMENT

IMAGE:4020395, mRNA, complete cds.
 BC031933
 BC031933.1 GI:21594465
 MGC.

Mus musculus (house mouse)

Mus musculus

Eukaryota; Euteleostomi;
 Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
 Strausberg, R.

1 (bases 1 to 2577)

Direct Submission

Submitted (06-JUN-2002) National Institutes of Health, Mammalian
 Gene Collection (MGC), Cancer Genomics Office, National Cancer
 Institute, 31 Center Drive, Room 11A03, Bethesda, MD 20892-2590,
 USA

NIH-MGC Project URL: <http://mgc.nci.nih.gov>

Contact: MGC help desk

Email: cgapbs-remail.nih.gov

Tissue Procurement: Gilbert Smith, Ph.D.

CDNA Library Preparation: Life Technologies, Inc.

CDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)

DNA Sequencing by: Baylor College of Medicine Human Genome

Sequencing Center

Center code: BCM-HGSC

Web site: <http://www.hgsc.bcm.tmc.edu/cdna/>

Contact: amg@bcm.tmc.edu

Gunaratne, P.H., Garcia, A.M., Lu, X., Hulyk, S.W., Hale, S.M.,

Yoon, V.S., Kowis, C.R., Lawrence, S., Martin, R.G., Muzny, D.M.,

Richards, S., Gibbs, R.A.

Clone distribution: MGC clone distribution information can be found

through the I.M.A.G.E. Consortium/LLNL at: <http://image.llnl.gov>

Series: IRAC Plate: 30 Row: K Column: 3

This clone was selected for full length sequencing because it

passed the following selection criteria: Hexamer frequency ORF

analysis, Similarity but not identity to protein.

Location/Qualifiers

1..2577

/organism="Mus musculus"

/mol_type="mRNA"

/strain="CZECH II"

/db_xref="taxon:10090"

/clone="MGC:25790 IMAGE:4020395"

/tissue_type="Mammary tumor metastasized to lung."

MMTV-LTR/Mnt1 model. Expression driven by an MMTV-LTR

enhancer."

/clone_lib="NCI CGAP_Lu30"

/lab_host="DH10B"

/note="Vector: pCMV-SPORT6"

112..1380

/codon_start=1

/product="RIKEN cdna 4632427C23 gene"

/protein_id="AAH31933.1"

/db_xref="GI:21594466"

/db_xref="LocusID:78753"

/translation="MSEILSRVTVSHRVEIWLILVAYLLQNVNSHLPKADPE

AFMYSEIHKHGVSEYEVATGYILSVNIPRGQTRLKGGSPVLLQRLGILG

DASWYINLNNLGFILADGFDVWMSNGTMSRKHKTLSIDQEFWAFSYDNA

BDLPVNFILQKQKVVYVSGTGMGTFATFMPDLAKHKIWMFALAPIATV

KYASPGTKFLIPDMIKVLFGQEFYQTRFQRLFIYLCGMILDQICSNILLL

GGFNTNNNRANVYVAHTPAGTSVQNLHWSQVNSGELRAFDMGSETKNQKCKQ

PTPIRYKVDMMVPTAMTGGDWLSNPDGDDVDDVLLSEVTLNLIYHKNIPWAHVDFIWG

LDAPQRYVEIHLMKQENLPQGTCTCKL"

BASE COUNT 790 a 496 c 528 g 763 t

ORIGIN

Query Match 48.9%; Score 1000.4; DB 10; Length 2577;

Best Local Similarity 83.8%; Pred. No. 1.2e-245;

Matches 1170; Conservative 0; Mismatches 216; Indels 10; Gaps 3;

QY 48 TTTTAAACCCCAAAATCTTCTTACTTAGAATAGTTGTACATTCGACGAGAA-AAAA 106

Db 33 TCTTCATCTCTGTTCTTCTTCTGACTCTAGATGTAGTTGTACATCTCTTACCAACAAAG 92

107 TAAATGACGATGTTGGACCATGTTGGAAACCTTGTGTCAGACAGTGGATGTTGTTCTCACAG 166
 Db CAAATGCCGATGTTGGACCATGTTGGAAACCTTGTGTCAGACAGTGGATGTTGTTCTCACAG 152
 Qy AATGGAATGTTGGATGTTGGATGTTGGATGTTGGATGTTGGATGTTGGATGTTGGATGTTGGAT 226
 Db AGTGGAGATGTTGGATGTTGGATGTTGGATGTTGGATGTTGGATGTTGGATGTTGGATGTTGGAT 212
 Qy ACATATGCCAATGTTGGATGTTGGATGTTGGATGTTGGATGTTGGATGTTGGATGTTGGATGTTGGAT 286
 Db ACATATGCCAATGTTGGATGTTGGATGTTGGATGTTGGATGTTGGATGTTGGATGTTGGATGTTGGAT 272
 Qy ACATCAAGGCTATCCCTGTTGAGGATGTTGAGGATGTTGAGGATGTTGAGGATGTTGAGGATGTTGAGGAT 346
 Db ACATCAAGGCTATCCCTGTTGAGGATGTTGAGGATGTTGAGGATGTTGAGGATGTTGAGGATGTTGAGGAT 332
 Qy TGTAAACAGGATGTTGAGGATGTTGAGGATGTTGAGGATGTTGAGGATGTTGAGGATGTTGAGGATGTTGAGGAT 406
 Db TGTAAACAGGATGTTGAGGATGTTGAGGATGTTGAGGATGTTGAGGATGTTGAGGATGTTGAGGATGTTGAGGAT 392
 Qy GTTACTGACGATGTTGAGGATGTTGAGGATGTTGAGGATGTTGAGGATGTTGAGGATGTTGAGGATGTTGAGGAT 466
 Db GTTACTGACGATGTTGAGGATGTTGAGGATGTTGAGGATGTTGAGGATGTTGAGGATGTTGAGGATGTTGAGGAT 452
 Qy TAGCTGGGCTTCAATCTGTCAGATGTTGAGGATGTTGAGGATGTTGAGGATGTTGAGGATGTTGAGGATGTTGAGGAT 526
 Db TAGCTGGGCTTCAATCTGTCAGATGTTGAGGATGTTGAGGATGTTGAGGATGTTGAGGATGTTGAGGATGTTGAGGAT 512
 Qy AAAGGCTGTTGTCAGGATGTTGAGGATGTTGAGGATGTTGAGGATGTTGAGGATGTTGAGGATGTTGAGGATGTTGAGGAT 586
 Db AAAGGCTGTTGTCAGGATGTTGAGGATGTTGAGGATGTTGAGGATGTTGAGGATGTTGAGGATGTTGAGGATGTTGAGGAT 572
 Qy CAGTTATGATGATGTTGAGGATGTTGAGGATGTTGAGGATGTTGAGGATGTTGAGGATGTTGAGGATGTTGAGGATGTTGAGGAT 646
 Db CAGTTATGATGATGTTGAGGATGTTGAGGATGTTGAGGATGTTGAGGATGTTGAGGATGTTGAGGATGTTGAGGATGTTGAGGAT 632
 Qy AACGGGCGAGGATGTTGAGGATGTTGAGGATGTTGAGGATGTTGAGGATGTTGAGGATGTTGAGGATGTTGAGGATGTTGAGGAT 706
 Db AACGGGCGAGGATGTTGAGGATGTTGAGGATGTTGAGGATGTTGAGGATGTTGAGGATGTTGAGGATGTTGAGGATGTTGAGGAT 692
 Qy TGCAATTTTCCAGATGTTGAGGATGTTGAGGATGTTGAGGATGTTGAGGATGTTGAGGATGTTGAGGATGTTGAGGATGTTGAGGAT 766
 Db TGCAATTTTCCAGATGTTGAGGATGTTGAGGATGTTGAGGATGTTGAGGATGTTGAGGATGTTGAGGATGTTGAGGATGTTGAGGAT 752
 Qy CATAGCCATGTTTAAAGCATGTTTAAAGCATGTTTAAAGCATGTTTAAAGCATGTTTAAAGCATGTTTAAAGCATGTTTAAAGCAT 826
 Db CATAGCCATGTTTAAAGCATGTTTAAAGCATGTTTAAAGCATGTTTAAAGCATGTTTAAAGCATGTTTAAAGCATGTTTAAAGCAT 812
 Qy GATGATCAAGGATGTTTAAAGCATGTTTAAAGCATGTTTAAAGCATGTTTAAAGCATGTTTAAAGCATGTTTAAAGCATGTTTAAAGCAT 886
 Db GATGATCAAGGATGTTTAAAGCATGTTTAAAGCATGTTTAAAGCATGTTTAAAGCATGTTTAAAGCATGTTTAAAGCATGTTTAAAGCAT 872
 Qy ACTTGTATTTTACCTTGTGTCAGGATGTTTAAAGCATGTTTAAAGCATGTTTAAAGCATGTTTAAAGCATGTTTAAAGCATGTTTAAAGCAT 946
 Db ACTTGTATTTTACCTTGTGTCAGGATGTTTAAAGCATGTTTAAAGCATGTTTAAAGCATGTTTAAAGCATGTTTAAAGCATGTTTAAAGCAT 932
 Qy ACTTCTGGGATGTTTAAAGCATGTTTAAAGCATGTTTAAAGCATGTTTAAAGCATGTTTAAAGCATGTTTAAAGCATGTTTAAAGCAT 1006
 Db ACTTCTGGGATGTTTAAAGCATGTTTAAAGCATGTTTAAAGCATGTTTAAAGCATGTTTAAAGCATGTTTAAAGCATGTTTAAAGCAT 992
 Qy CCACATCTTGTGTCAGGATGTTTAAAGCATGTTTAAAGCATGTTTAAAGCATGTTTAAAGCATGTTTAAAGCATGTTTAAAGCATGTTTAAAGCAT 1066
 Db CCACATCTTGTGTCAGGATGTTTAAAGCATGTTTAAAGCATGTTTAAAGCATGTTTAAAGCATGTTTAAAGCATGTTTAAAGCATGTTTAAAGCAT 1052
 Qy TGTGTAATCTGGGATGTTTAAAGCATGTTTAAAGCATGTTTAAAGCATGTTTAAAGCATGTTTAAAGCATGTTTAAAGCATGTTTAAAGCAT 1126
 Db TGTGTAATCTGGGATGTTTAAAGCATGTTTAAAGCATGTTTAAAGCATGTTTAAAGCATGTTTAAAGCATGTTTAAAGCATGTTTAAAGCAT 1112
 Qy GCCAATCTTGTGTCAGGATGTTTAAAGCATGTTTAAAGCATGTTTAAAGCATGTTTAAAGCATGTTTAAAGCATGTTTAAAGCATGTTTAAAGCAT 1186
 Db GCCAATCTTGTGTCAGGATGTTTAAAGCATGTTTAAAGCATGTTTAAAGCATGTTTAAAGCATGTTTAAAGCATGTTTAAAGCATGTTTAAAGCAT 1172

1187 AGCTCAGGATGTTTAAAGCATGTTTAAAGCATGTTTAAAGCATGTTTAAAGCATGTTTAAAGCATGTTTAAAGCATGTTTAAAGCAT 1246
 Db AGCTCAGGATGTTTAAAGCATGTTTAAAGCATGTTTAAAGCATGTTTAAAGCATGTTTAAAGCATGTTTAAAGCATGTTTAAAGCAT 1232
 Qy CCTCATCTACCATGTTTAAAGCATGTTTAAAGCATGTTTAAAGCATGTTTAAAGCATGTTTAAAGCATGTTTAAAGCATGTTTAAAGCAT 1306
 Db CCTCATCTACCATGTTTAAAGCATGTTTAAAGCATGTTTAAAGCATGTTTAAAGCATGTTTAAAGCATGTTTAAAGCATGTTTAAAGCAT 1292
 Qy TGCTCTCAGGATGTTTAAAGCATGTTTAAAGCATGTTTAAAGCATGTTTAAAGCATGTTTAAAGCATGTTTAAAGCATGTTTAAAGCAT 1366
 Db TGCTCTCAGGATGTTTAAAGCATGTTTAAAGCATGTTTAAAGCATGTTTAAAGCATGTTTAAAGCATGTTTAAAGCATGTTTAAAGCAT 1349
 Qy TTCCAGGATGTTTAAAGCATGTTTAAAGCATGTTTAAAGCATGTTTAAAGCATGTTTAAAGCATGTTTAAAGCATGTTTAAAGCATGTTTAAAGCAT 1426
 Db TTCCAGGATGTTTAAAGCATGTTTAAAGCATGTTTAAAGCATGTTTAAAGCATGTTTAAAGCATGTTTAAAGCATGTTTAAAGCATGTTTAAAGCAT 1403
 Qy CCTCTCAGGATGTTTAAAGCATGTTTAAAGCATGTTTAAAGCATGTTTAAAGCATGTTTAAAGCATGTTTAAAGCATGTTTAAAGCATGTTTAAAGCAT 1442
 Db CCTCTCAGGATGTTTAAAGCATGTTTAAAGCATGTTTAAAGCATGTTTAAAGCATGTTTAAAGCATGTTTAAAGCATGTTTAAAGCATGTTTAAAGCAT 1419

RESULT 4
 AL353113 43456 bp DNA linear PRI 20-SEP-2001
 LOCUS Human DNA sequence from clone RPL1-30415 on chromosome 10. Contains
 DEFINITION the gene for a novel lipase, the gene for a novel protein similar
 to ankyrin, ESTs, STSs and GSSs, complete sequence.
 ACCESSION AL353113
 VERSION AL353113.7 GI:15130723
 KEYWORDS HTG; ankyrin; lipase.
 SOURCE Homo sapiens (human)
 ORGANISM Homo sapiens
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
 1 (bases 1 to 43456)
 Direct Submission
 Submitted (23-AUG-2001) Sanger Centre, Hinxton, Cambridgeshire,
 CB10 1SA, UK. E-mail enquiries: humquery@sanger.ac.uk
 Requests: clonerequest@sanger.ac.uk
 On Aug 9, 2001 this sequence version replaced gi:14669221.
 During sequence assembly data is compared from overlapping clones.
 Where differences are found these are annotated as variations
 together with a note of the overlapping clone name. Note that the
 variation annotation may not be found in the sequence submission
 corresponding to the overlapping clone, as we submit sequences with
 only a small overlap as described above.
 This sequence was finished as follows unless otherwise noted: all
 regions were either double-stranded or sequenced with an alternate
 chemistry or covered by high quality data (i.e. phred quality >= 30); an attempt was made to resolve all sequencing problems, such
 as compressions and repeats; all regions were covered by at least
 one plasmid subclone or more than one M13 subclone; and the
 assembly was confirmed by restriction digest. The following
 abbreviations are used to associate primary accession numbers given
 in the feature table with their source databases: Em: EMBL; Sw:
 SWISSPROT; Tr: TREMBL; Wp: WORMPEP; Information on the WORMPEP
 database can be found at
 http://www.sanger.ac.uk/Projects/C_elegans/wormpep This sequence
 was generated from part of bacterial clone contigs of human
 chromosome 10, constructed by the Sanger Centre Chromosome 10
 Mapping Group. Further information can be found at
 http://www.sanger.ac.uk/HGP/Chr10
 RPL1-30415 is from the library RPL1-11.2 constructed by the group
 of Pieter de Jong. For further details see
 http://www.chori.org/bacpac/home.htm
 VECTOR: pBAC3.6
 IMPORTANT: This sequence is not the entire insert of clone
 RPL1-30415. It may be shorter because we sequence overlapping
 sections only once, except for a short overlap.
 The true left end of clone RPL1-399019 is at 41457 in this
 sequence. The true right end of clone RPL1-186014 is at 2000 in
 this sequence.

* 28737 46598: contig of 17862 bp in length
 * 46599 gap of 100 bp
 * 46699 71912: contig of 25214 bp in length
 * 71913 gap of 100 bp
 * 72013 96612: contig of 24600 bp in length
 * 96613 gap of 100 bp
 * 96713 128331: contig of 31619 bp in length
 * 128332 128431: gap of 100 bp
 * 128432 158195: contig of 29764 bp in length.

FEATURES

source
 1..158195
 /organism="Homo sapiens"
 /mol_type="genomic DNA"
 /db_xref="taxon:9606"
 /clone="RP11-115N17"
 /clone.lib="RPC1-11 Human Male BAC"
 1..1002
 /note="assembly_fragment"
 1103..2388
 /note="assembly_fragment"
 2489..3525
 /note="assembly_fragment"
 3626..16527
 /note="assembly_fragment"
 16628..28636
 /note="assembly_fragment"
 clone_end:77
 vector_side:right
 28737..46598
 /note="assembly_fragment"
 46699..71912
 /note="assembly_fragment"
 72013..96612
 /note="assembly_fragment"
 96713..128331
 /note="assembly_fragment"
 128432..158195
 /note="assembly_fragment"
 clone_end:876
 vector_side:left
 BASE COUNT 48628 a 30917 c 30188 g 47557 t 905 others
 ORIGIN

Query Match 43.6%; Score 890.4; DB 2; Length 158195;

Best Local Similarity 99.9%; Pred. No. 2.8e-217; Mismatches 1; Indels 0; Gaps 0;

Matches 891; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1126 AGCCAACTCCTCTAAGGTACAGAGTCAGAGATATGACGGTCCCTACAGCAATGTGGACAG 1185

DB 19482 AGCCAACTCCTCTAAGGTACAGAGTCAGAGATATGACGGTCCCTACAGCAATGTGGACAG 19423

QY 1186 GAGGTGAGGACTGGCTTTCAATCCAGAGAGCGTGAATGCTGCTCTCTGAGGTGACCA 1245

DB 19422 GAGGTGAGGACTGGCTTTCAATCCAGAGAGCGTGAATGCTGCTCTCTGAGGTGACCA 19363

QY 1246 ACCTCATCTACCATAGATATTCCTGAATGGCTCAGTGATTTTCATCTCGGGTTGG 1305

DB 19362 ACCTCATCTACCATAGATATTCCTGAATGGCTCAGTGATTTTCATCTCGGGTTGG 19303

QY 1306 ATGCTCCTCACCGTATGACCAATCAATCAATCATCTGATGATGAGGAGGAGACCAACC 1365

DB 19302 ATGCTCCTCACCGTATGACCAATCAATCAATCATCTGATGATGAGGAGGAGACCAACC 19243

QY 1366 TTTCACGAGGAGGTTGAGGCGGTATTTGTGAGCATCTGACACTGACGATCTTAGGACA 1425

DB 19242 TTTCACGAGGAGGTTGAGGCGGTATTTGTGAGCATCTGACACTGACGATCTTAGGACA 19183

QY 1426 ACCTCCTGAGGATGGGGCTAGGACCCATGAAGGAGCAATTCAGGAGGAGGAGACCTAG 1485

DB 19182 ACCTCCTGAGGATGGGGCTAGGACCCATGAAGGAGCAATTCAGGAGGAGGAGACCTAG 19123

QY 1486 TATACATTTTTCAGATTCCTGCACTTGGCACTTAATCCGACACTTACATTTT 1545

DB 19122 TATACATTTTTCAGATTCCTGCACTTAATCCGACACTTACATTTTACATTTT 19063

QY 1546 TTCTCTAAATTAAGTACTTATTAGGTAAATAGAGGTTTGTATGCTATATATATTTCT 1605

DB 19062 TTCTCTAAATTAAGTACTTATTAGGTAAATAGAGGTTTGTATGCTATATATATTTCT 19003

QY 1606 ACCATCTTCAAGGTTAGGTTTACCTGTAGTACGAGCAAAATATCTAGACATTTCTATATC 1665

DB 19002 ACCATCTTCAAGGTTAGGTTTACCTGTAGTACGAGCAAAATATCTAGACATTTCTATATC 18943

QY 1666 ATTCAAGTAAATCTCTTTAAACACCTATTTTCTATTAAGCCATATTTTGGAGCA 1725

DB 18942 ATTCAAGTAAATCTCTTTAAACACCTATTTTCTATTAAGCCATATTTTGGAGCA 18883

QY 1726 CTAAAGTAAATGCGCAAAATTTGGACAGATATTGAGTCTGAGTCTGTGGATTATTGTG 1785

DB 18882 CTAAAGTAAATGCGCAAAATTTGGACAGATATTGAGTCTGTGGATTATTGTG 18823

QY 1786 ACTTTGACAAAATAAGCTAGACATTTTCACTTTGTTGCCACAGAGACATAACACCTACCTC 1845

DB 18822 ACTTTGACAAAATAAGCTAGACATTTTCACTTTGTTGCCACAGAGACATAACACCTACCTC 18763

QY 1846 AGGAAGTCTGAGTCTTTAAGGACACACACAAAATCAGTTTACAGTATGGAGAAA 1905

DB 18762 AGGAAGTCTGAGTCTTTAAGGACACACACAAAATCAGTTTACAGTATGGAGAAA 18703

QY 1906 TCTATGTTAAGCATTTCTCAGAAATAGGCCAAAGTTTATAGTTGTCATCTCAGGGAAGAAA 1965

DB 18702 TCTATGTTAAGCATTTCTCAGAAATAGGCCAAAGTTTATAGTTGTCATCTCAGGGAAGAAA 18643

QY 1966 TTTTATAGGATGTTAAGGTTCTCCTCAATAAATGCAATTTGCAATTACATAAA 2017

DB 18642 TTTTATAGGATGTTAAGGTTCTCCTCAATAAATGCAATTTGCAATTACATAAA 18591

RESULT 6

AXS21801

LOCUS

DEFINITION

ACCESSION

VERSION

KEYWORDS

SOURCE

ORGANISM

REFERENCE

AUTHORS

TITLE

JOURNAL

FEATURES

source

1..1116

/organism="Homo sapiens"

/mol_type="genomic DNA"

/db_xref="taxon:9606"

BASE COUNT 326 a 234 c 245 g 311 t

ORIGIN

Query Match 36.8%; Score 752; DB 6; Length 1116;

Best Local Similarity 97.4%; Pred. No. 7.1e-182;

Matches 783; Conservative 0; Mismatches 0; Indels 21; Gaps 1;

QY 273 AGTGAATCATCCAAATCAAGCTATCCCTGTGAGGAATATGAGTGCACACTGAAGT 332

DB 73 AGTGAATCATCCAAATCAAGCTATCCCTGTGAGGAATATGAGTGCACACTGAAGT 132

QY 333 GGGTATATCTTTCTCTTAAACAGGATTCCTCAGGCGCTAGTGCACACCTAAAGACAGGT 392

DB 133 GGGTATATCTTTCTCTTAAACAGGATTCCTCAGGCGCTAGTGCACACCTAAAGACAGGT 192

QY 393 TCAGGCGCTGTGGTGTACTGACGATGCGCTAGTTGAGGAGTGTGACACTGGAATTC 452

Db 193 TCAGGCGCTGTGTGTTACTGTCAGCATGCGCTAGTTGGAGGTGCTAGCACTGGATTTC 252

Qy 453 AACCTGCCCAACAATAGCTGGCTGCTTCATCTGGCAGATGCTGGTTTGAAGTGGATG 512

Db 253 AACCTGCCCAACAATAGCTGGCTGCTTCATCTGGCAGATGCTGGTTTGAAGTGGATG 312

Qy 513 GGAACAGCAGGAGGAAACCGCTGCTTCAAAAACACAGACACTCTCCATAGACCAAGAT 572

Db 313 GGAACAGCAGGAGGAAACCGCTGCTTCAAAAACACAGACACTCTCCATAGACCAAGAT 372

Qy 573 GAGTTCGGGCTTTCAGTTATAGATAGATGAGTGGCTTGAAGTTCCTGCTGAGTGAATAAC 632

Db 373 GAGTTCGGGCTTTCAGTTATAGATAGATGAGTGGCTTGAAGTTCCTGCTGAGTGAATAAC 432

Qy 633 TTATTTTTCAGAAAAACGGCCAGGAAAGATCTATTATGTCGGCTATTTCACAGGCGACC 692

Db 433 TTATTTTTCAGAAAAACGGCCAGGAAAGATCTATTATGTCGGCTATTTCACAGGCGACC 492

Qy 693 ACCATGGGCTTTATTCATTTTCCACCATGCGCAGAGCTGGCTCAGAAAAATCAAAATGTAT 752

Db 493 ACCATGGGCTTTATTCATTTTCCACCATGCGCAGAGCTGGCTCAGAAAAATCAAAATGTAT 552

Qy 753 TTGCTTTTACCCATGACCTGTTAGCATGCAAAAGCCCGGACCAAAATTTTG 812

Db 553 TTGCTTTTACCCATGACCTGTTAGCATGCAAAAGCCCGGACCAAAATTTTG 612

Qy 813 TTGCTCCAGATATGATGATCAAGGATTTTGGCAAAAAGAAATTTCTGTATCAGACC 872

Db 613 TTGCTCCAGATATGATGATCAAGGATTTTGGCAAAAAGAAATTTCTGTATCAGACC 672

Qy 873 AGATTTCTCAGACACTGTTATTACCTTTGTCGGCAGGTGATCTTGTATCAGATTGT 932

Db 673 AGATTTCTCAGACACTGTTATTACCTTTGTCGGCAGGTGATCTTGTATCAGATTGT 732

Qy 933 AGTAATATCATGTTACTTCTGGTGGATTCAACCAACAATATGAACAT 982

Db 733 AGTAATATCATGTTACTTCTGGTGGATTCAACCAACAATATGAACATGATCTCAT 792

Qy 983 -----GAGCGAGCAAGTGTATATGTCGCCCACTCTTGTGCGAAACATCTGTG 1031

Db 793 GGTGTTGTTACAGAGCCGAGCATGTATATGTCGCCCACTCTTGTGCGAAACATCTGTG 852

Qy 1032 CAAAATATTCACACTGGAGCCAG 1055

Db 853 CAAAATATTCACACTGGAGCCAG 876

RESULT 7

AC015506/c 76653 bp DNA linear HTG 13-JUL-2000

LOCUS Homo sapiens clone RP11-21124, LOW-PASS SEQUENCE SAMPLING.

DEFINITION AC015506

ACCESSION AC015506.3 GI:9124007

VERSION HTG; HTGS PHASE0.

KEYWORDS Homo sapiens (human)

SOURCE Homo sapiens

ORGANISM Homo sapiens

REFERENCE 1 (bases 1 to 76653)

AUTHORS Birren,B., Linton,L., Nuebaum,C. and Lander,E.

TITLE Homo sapiens, clone RP11-21124

JOURNAL Unpublished

REFERENCE 2 (bases 1 to 76653)

AUTHORS Birren,B., Linton,L., Nuebaum,C., Lander,E., Allen,N., Anderson,M., Baldwin,J., Barna,N., Beckerly,R., Boguslavsky,L., Boukhgalter,B., Brown,A., Castle,A., Colangelo,M., Collins,S., Collymore,A., Cooke,P., DeArelano,K., Dewar,K., Domino,M., Donelan,L., Doyle,M., Ferreira,P., FitzHugh,W., Forrest,C., Funke,R., Gage,D., Galagan,J., Gardyna,S., Grant,G., Hagos,B., Heaford,A., Horton,L., Howland,J.C., Johnson,R., Jones,C., Kann,L., Karatas,A., Klein,J., Lehoszky,J., Lieu,C., Locke,K., Macdonald,P., Marquis,N., McEwan,P., McGurk,A., McKernan,K., McLaughlin,J., Meldrum,J., Morrow,J., Naylor,J., Norman,C.H., O'Connor,T., O'Donnell,P.,

Peterson,K., Pollara,V., Riley,R., Roy,A., Santos,R., Severy,P., Stange-Thomann,N., Stojanovic,N., Subramanian,A., Talamas,J., Tesfaye,S., Tirrell,A., Vassiliev,H., Vo,A., Wheeler,J., Wu,X., Wyman,D., Ye,W.J., Zimmer,A. and Zody,M.

Direct Submission

Submitted (16-NOV-1999) Whitehead Institute/MIT Center for Genome Research, 320 Charles Street, Cambridge, MA 02141, USA

On Jul 13, 2000 this sequence version replaced gi:6468098.

All repeats were identified using RepeatMasker:

Smit, A.F.A. & Green, P. (1996-1997)

http://ftp.genome.washington.edu/RW/RepeatMasker.html

----- Genome Center

Center: Whitehead Institute/ MIT Center for Genome Research

Center code: WIBR

Web site: http://www-seq.wi.mit.edu

Contact: sequence_submissions@genome.wi.mit.edu

----- Project Information

Center project name: L4028

Center clone name: 21_I_24

* NOTE: This record contains 87 individual

* sequencing reads that have not been assembled into

* contigs. Runs of N are used to separate the reads

* and the order in which they appear is completely

* arbitrary. Low-pass sequence sampling is useful for

* identifying clones that may be gene-rich and allows

* overlap relationships among clones to be deduced.

* However, it should not be assumed that this clone

* will be sequenced to completion. In the event that

* the record is updated, the accession number will

* be preserved.

1 795: contig of 795 bp in length

* 796 895: gap of 100 bp

* 896 1656: contig of 761 bp in length

* 1657 1756: gap of 100 bp

* 1757 2558: contig of 802 bp in length

* 2559 2658: gap of 100 bp

* 2659 3434: contig of 776 bp in length

* 3435 3534: gap of 100 bp

* 3535 4301: contig of 767 bp in length

* 4302 4401: gap of 100 bp

* 4402 5199: contig of 798 bp in length

* 5200 5299: gap of 100 bp

* 5300 6085: contig of 786 bp in length

* 6086 6185: gap of 100 bp

* 6186 6972: contig of 787 bp in length

* 6973 7072: gap of 100 bp

* 7073 7844: contig of 772 bp in length

* 7845 7944: gap of 100 bp

* 7945 8731: contig of 787 bp in length

* 8732 8831: gap of 100 bp

* 8832 9625: contig of 794 bp in length

* 9626 9725: gap of 100 bp

* 9726 10519: contig of 794 bp in length

* 10520 10619: gap of 100 bp

* 10620 11417: contig of 798 bp in length

* 11418 11517: gap of 100 bp

* 11518 12297: contig of 779 bp in length

* 12297 12396: gap of 100 bp

* 12397 13180: contig of 784 bp in length

* 13181 13280: gap of 100 bp

* 13281 14065: contig of 785 bp in length

* 14066 14165: gap of 100 bp

* 14166 14949: contig of 784 bp in length

* 14950 15049: gap of 100 bp

* 15050 15822: contig of 773 bp in length

* 15823 15922: gap of 100 bp

* 15923 16718: contig of 796 bp in length

* 16719 17621: contig of 803 bp in length

* 17622 17721: gap of 100 bp

* 17722 18517: contig of 796 bp in length

* 18518 18617: gap of 100 bp

[illegible]


```
RESULT 9
AX411234
LOCUS AX411234 2481 bp DNA linear PAT 14-JUN-2002
DEFINITION Sequence 3881 from Patent WO0229103.
ACCESSION AX411234
VERSION AX411234.1 GI:21443939
KEYWORDS
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE
1. Alvares, C., Horne, D., Peres-da-Silva, S. and Vockley, J.G.
Gene expression profiles in liver cancer
Patent: WO 0229103-A 3881 11-APR-2002;
GENE LOGIC INC (US)
FEATURES
source
1. .2481
/organism="Homo sapiens"
/mol_type="genomic DNA"
/db_xref="taxon:9606"
/note="EMBL/GenBank Accession No. Z31690"
BASE COUNT 692 a 477 c 513 g 799 t
ORIGIN
Query Match 25.8%; Score 527.8; DB 6; Length 2481;
Best Local Similarity 66.6%; Pred. No. 2.7e-124;
Matches 751; Conservative 0; Mismatches 376; Indels 0; Gaps 0;
QY 240 AAGCTGTGGACCCAGAGAGCTTATGAATATTAGTGAATATCCACATCAAGGCTAT 299
Db |||||
QY 112 ACAGCTGTGGATCCTGAAACAAACATGAATGTGAGTGAATATCTCTTACTGGGATTC 171
Db |||||
QY 300 CCTGTGAGGAATATGAAGTGCACACTCAAGATGGGTATATCTCTCTGTTAAACAGATT 359
Db |||||
QY 172 CCTAGTAGGAATACCTAGTTGAGACAGAGATGAATATCTGTGCTTAAACCGAATT 231
Db |||||
QY 360 COTGAGGCCCTAGTGCAACCTAAGAGACAGGTTCCAGGCTGTGGTGTACTGCAGCAT 419
Db |||||
QY 232 COTCATGGGAGGAGAACCATCTTGACAAAGGTCCCAACAGTGTCTCTCTCCCAACAT 291
Db |||||
QY 420 GGCCTAGTTGGAGGCTGACCACTGATTTCCACCTGCCCAACATAGCTGGGCTTC 479
Db |||||
QY 292 GGCCTGCTGGCAGATCTCTAGTAACCTGGGTCAAACTGTCACACAGAGCTGGGCTTC 351
Db |||||
QY 480 ATTCTGGCAGATGCTGGTGTGATGGGAAACAGACAGGGAACAGCGCTGTCT 539
Db |||||
QY 352 ATCTCTGTGATGCTGGTGTGATGGGAAACAGACAGGGAATACCTGTCT 411
Db |||||
QY 540 CGAAACACAGACACTCTCCATAGACCAAGATGAGTTCTGGGCTTTCAGTTATGATGAG 599
Db |||||
QY 412 CGAAACACATAAGACACTCTCAGTTTCTCAGGATGAATTTCTGGGCTTTCAGTTATGATGAG 471
Db |||||
QY 600 ATGGCTAGTTGACCTTCTGTCAGTATAAATTTATTTTGCAGAAACAGGCGCCAGAA 659
Db |||||
QY 472 ATGGCAAAATATGACTACAGCTTCCATTAATCTCAATTAATAAATGCGCAAGAA 531
Db |||||
QY 660 AAGATCTATTATGTCGGCTATTACAGGGCCACCAATGCGCTTTATTCGATTTTCCACC 719
Db |||||
QY 532 CAAGTGTATTATGTTGGTCAATCTCAAGGCCACCACTATAGTGTTTATAGCATTTTCACAG 591
Db |||||
QY 720 ATGCCAGAGCTGGCTCAGAAATCAAAATGATTTTGTGTTAGCACCACATAGCCACTGTT 779
Db |||||
QY 592 ATCCCTGAGCTGGCTTAAAGGATTAATTTTGTGCTGGTCTGCTGGCTTCCGCTC 651
Db |||||
QY 780 AAGCATGCAAAAGCCCGGACCAAAATTTTGTGCTGCCAGATATGATGATCAAGGGA 839
Db |||||
QY 652 GCCTTCTGTACTAGCCCTATGGCAATTAAGGACGATTACCATCATCTCATTAAGGAC 711
Db |||||
QY 840 TGTGTTGGCAAAAGAAATTTCTGATCAGACAGATTTCTCAGACAACTGTTATTATAC 899
Db |||||

Db 712 TTATTGGAGACAAAGAAATTTCTCCCGAGAGTGGTGTTCGAGTGGCTGGTACCCAC 771
QY 900 CTTTGTGGCCAGGTGATTTCTTGATCAGATTTGTAGTAATATCATGTTACTTCTTGGGTGGA 959
Db |||||
QY 772 GTTTCAGCTCATGTCTATCTGAAGGAGCTCTGTGAAATCTCTGTTTCTTCTGTGTGGA 831
Db |||||
QY 960 TTCAACACCAACAAATATGAACATGAGCCGAGCAAGTGTATATCTGTCGCCACACTCTTCT 1019
Db |||||
QY 832 TTATATGAGAGAAATTTAAATATGTCTAGAGTGGATATATACACACATTTCTCTCTCT 891
Db |||||
QY 1020 GGAACATCTGTGCAAAATATTTTACACTGGAGCCAGGAGTGAATTTCTGGTGAATCTCCG 1079
Db |||||
QY 892 GGAACCTTCTGTGCAAAACATGTTACACTGGAGCCAGGCTGTTAAATTTCAAAAGTTTCAA 951
Db |||||
QY 1080 GCATTTGACTGGGGAGTGAGACCAAAATCTGGAATAATGCAATCAGCAACTCTCTGTA 1139
Db |||||
QY 952 GCCTTTGACTGGGAGAGCAGTCCCAAGAAATATTTTCAATACACAGAGTATCTCTCC 1011
Db |||||
QY 1140 AGGTACAGAGTCCAGAGATATGACGGTCCCTACAGCAATGTGGACAGGAGTCCAGACTGG 1199
Db |||||
QY 1012 ACATACATGTGAAGGACATGCTTGTGCGACTGCACTGTGGAGCGGGGTCCAGACTGG 1071
Db |||||
QY 1200 CTTTCAATCCAGAGAGCTGAAATGCTCTCTGAGGTGACCAACTCATCTACCAT 1259
Db |||||
QY 1072 CTTGCAGATGTCTACGAGTCAATATCTTACTGACTCAGATCACCACACTTGGTGTCCAT 1131
Db |||||
QY 1260 AAGATATTTCTGAATGGGCTCACGCTGATTTTCATCTGGGTTTGGATGCTCTCACCGT 1319
Db |||||
QY 1132 GAGAGCATTTCCGAAATGGAGCATCTTGACTTTCATTTGGGCTGATGCTCCCTTGGAGG 1191
Db |||||
QY 1320 ATGTACATGAATATCATCTCATCTGATGACAGCAGGAGGAGACCAACT 1366
Db |||||
QY 1192 CTTTAAATTAATTAATTAATCTTAATGAGAAATATCAGTGAAGCT 1238
Db |||||

RESULT 10
HSLVACLY
LOCUS HSLVACLY 2481 bp mRNA linear PRI 17-FEB-1997
DEFINITION H. sapiens (HepG2) LAL mRNA for lysosomal acid lipase.
ACCESSION Z31690
VERSION Z31690.1 GI:506430
KEYWORDS LAL; lipase; lysosomal acid lipase.
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE
1. Du, H. and Gregory, G.A.
Structural Conservation of Putative Functional Motifs between Mouse
and Human Lysosomal Acid Lipase
Unpublished
2. (bases 1 to 2481)
Du, H.
Direct Submission
Submitted (05-APR-1994) Hong Du, Division of Human Genetics,
Children's Hospital Medical Center, 3333 Burnet Street, Cincinnati,
OH, 45229-3032, USA
FEATURES
source
1. .2481
/organism="Homo sapiens"
/mol_type="mRNA"
/strain="HepG2"
/db_xref="taxon:9606"
/clone="clone hLAL"
/cell_type="hepatoma"
/tissue_type="liver"
/clone_lib="human liver HepG2 cDNA library/lambda ZAP"
1. .2481
/gene="LAL"
34.1233
/gene="LAL"
/citation=[1]
/codon_start=1
gene
CDS
```

```

/product="lysosomal acid lipase"
/protein_id="CAA83495.1"
/db_xref="GI:506431"
/translation="MKQRFGLVLCVLTWLSHSGRGLTAVDPETNMVSEIISYW
GPFSEYLVETEDGYILCLNRIPIHGRKNHSDKPKPVVPLQHLGLADSSNWNVLANS
SLGFIILADAGFVDMGNSRGNTWSRKHKTLVSQDFWAFSYDENAKYDLPASINFL
NKTGQSVYVGHSGCTTIGFIAPSOIPELAKRIKMFALGPVAVFCTSPMAKLR
LPDLKIDLPGEKFLPQSAFLKWLGHVCHVILKELGNCILLCGFENRNLNMSR
VDYTHSPAGTSVQNLHWSQAVFKQFADFWSGSSAKNYEHYNQSPPTNYVDML
VPTAVWGGHDMADYVDVNIILLTOITNLVFEHSIPWEHEDFINGLDAPRLNKKII
NLMRYQ"
34..96
/gene="LAL"
/citation=[1]
/mat_peptide
97..1230
/product="lysosomal acid lipase"
/gene="LAL"
/citation=[1]
79
/gene="LAL"
/citations=[1]
/replace="c"
100
/gene="LAL"
/citations=[1]
/replace="g"
2459..2464
/gene="LAL"
2480
/gene="LAL"
692 a 477 c 513 g 799 t
BASE COUNT
ORIGIN
Query Match 25.8%; Score 527.8; DB 9; Length 2481;
Best Local Similarity 66.6%; Pred. No. 2.7e-124;
Matches 751; Conservative 0; Mismatches 376; Indels 0; Gaps 0;
QY 240 AAAGCTGTGGACCCAGAGCAATTCATCAATATTAGTGAATCATCCAAATCAAGGCTAT 299
DB 112 ACAGCTGTGGATCTCGAAACAAACATGAATGTGAGTGAATATCTCTACTGGGATTTC 171
QY 300 CCCTGTGAGGAATATGAAGTCGCACTGAAGTGGGTATATCTCTTCTGTAAACAGATT 359
DB 172 CCTAGTAGGAATACCTAGTTGAGACAGAGATGATATCTCTGCTTAAACCAAT 231
QY 360 CCTGAGGCTTGTGCACTTAAAGACAGAGTTCAGGCTGTGGTGTGTTACTGCAAGCAT 419
DB 232 CCTCATGGGAGGAGAACCATTCGACAAAGGTCCCAACAGTGTCTTCTCTGCAACAT 291
QY 420 GGCCTAGTTGAGGTGTGACCACTGATGATGATGATGATGATGATGATGATGATGATG 479
DB 292 GGCCTGTGGCAGATTCCTAGTAACTGGGTCAAAACCTTGGCAACAGACGCTGGGCTTC 351
QY 480 ATTCTGGCAGATGCTGCTTTTTCAGCTGTGATGATGATGATGATGATGATGATGATG 539
DB 352 ATTCTGTGATGCTGCTTTTTCAGCTGTGATGATGATGATGATGATGATGATGATG 411
QY 540 CGAAACACAGACACTCTCATAGACCAAGATGATGATGATGATGATGATGATGATGATG 599
DB 412 CGAAACACATAGACACTCTCATAGATGATGATGATGATGATGATGATGATGATGATG 471
QY 600 ATGGCTAGGTTTTCAGCTTCTGAGTGAATGATGATGATGATGATGATGATGATGATG 659
DB 472 ATGGCAAAATATGACCTACAGCTTCCATTAACCTTCAATCTGATATAAATCTGSCA 531
QY 660 AAGATCTATTATGTCGGTATTCACAGGCAACCAATGAGGCTTTATGATGATTTCCACC 719
DB 532 CAAGTGATATTATGTCGGTCAATCTCAAGGCAACCAATAGGCTTTTATGATGATTTTCA 591
QY 720 ATGCCAGAGCTGGTTCAGAAATCAAAATGATGATGATGATGATGATGATGATGATG 779
DB 592 ATCCCTGAGCTGGCTAAAGAGGATTAATGATGATGATGATGATGATGATGATGATG 651

```

```

839 AACATGCAAAAGCCCGGAGCAAAATTTTCTGTCGCATATGATGATCAAGGGA
711 GCCTTCTGTACTAGCCCTTATGGCAATTAAGGAGATTACAGATCATCTCATTAAGGAC
899 TTGTTTGGCAAAAGAAATTTCTGTATCAGACCAAGATTTCTCAGACAACTTGTATTAC
771 TTAATTTGAGACAAAGAAATTTCTCCCAAGAGTGGCTTTTGAAGTGGTGGTACCCAC
959 CTTTGTGGCCAGGTGATTTCTGTATCAGATTTGTAGTAATATCATGTTACTTCTGGTGGGA
831 GTTTGCACTCATGTACATGAGAGCTCTGTGGAATCTCTGTTTCTTCTGTGGGA
960 TTCAACCAACCAATATGAACATGAGCCGAGCAAGTGTATATCTGCCCACTCTTGTCT
891 TTTAATGAGAGAAATTTAAATATGTCTAGAGTGGATGATATATACACACATTTCTCTGCT
1079 GGAACATCTGTGCAAAATATTTCTACATCTGAGCCAGCAGTGAATCTGCTGAGCACTCCG
951 GGAACCTTCTGTGCAAAACATGTTACATCTGGAGCAGCTGTAAATTTCCAAAGTTTCAA
1139 GCATTTGACTGGGGAGTGAGACCAAAATCTGGAAAAATGCAATCAGCAACTCTCTGTGA
1011 GCTTTGACTGGGAGCAGTGCAGAAATTTATTTTCAACCAAGAGTATCTCTCTCC
1199 AGGTACAGAGTCAGAGATATGACGCTCCCTACAGCAATGTGGACAGAGTCAAGAGTGG
1071 ACATACAAATGTGAAGGACATGCTTTGTGCGCACTSCAGTCTGGAGCGGGGTCAAGACTGG
1259 CTTTCAAAATCCAGAGAGCAGTGAAATGCTCTCTGAGGTGACCAACCTCATCTACCAT
1131 CTTGCAATGTCTACGAGCTCAATATCTTACTGACTCAGATCACCACCTTGGTGTCCAT
1319 AAGAAATTTCTGAATGGCTCAGCTGATTTCACTGGGGTTTGGATGCTCTCCACCGT
1191 CAGAGCATTCGGAATGGGAGCATCTTCACTTCAATTTGGGGCCTGGATGCTTGGAGG
1366 ATGTACATGAATCAATCCATCTCATGACGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG
1238 CTTTATAATAAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAAT

```

```

RESULT 11
HSU08464
LOCUS Human lysosomal acid lipase mRNA, complete cds.
DEFINITION Human lysosomal acid lipase mRNA, complete cds.
ACCESSION U08464
VERSION U08464.1
KEYWORDS GI:505052
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE 1 (bases 1 to 2481)
AUTHORS Du,H. and Gregory,G.A.
TITLE Structural Conservation of Putative Functional Motifs between Mouse
and Human Lysosomal Acid Lipase
JOURNAL Unpublished
REFERENCE 2 (bases 1 to 2481)
AUTHORS Du,H.
TITLE Direct Submission
JOURNAL Submitted (08-APR-1994) Hong Du, Division of Human Genetics,
Children's Hospital Medical Center, 3333 Bernet Street, Cincinnati,
OH 45229-3039, USA
FEATURES
Location/Qualifiers
1..2481
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/clone="clone hLAL"
/cell_line="HepG2"
/cell_type="hepatoma"
/tissue_type="liver"

```


/clone_lib="HepG2 cDNA library/Lambda ZAP"

5'UTR
CDS

```
1. .33_1233
34. .1233
/codon_start=1
/product="lysosomal acid lipase"
/protein_id="AAB60328.1"
/db_xref="GI:505053"
/translaton="MKMRFLQLVCLVLTWLTSHSGSRGKLTAVDPETNMNVSRIISVW
GFPSEVLVEDDCYIILCLNRIPEGRKNHSDKGPVPLQHGILLADSSNWTNLANS
SLKTLADAGDVMGMHSGRNTWSRKHTLSVQDEFWAFSYDEMAYKIDLPASINFIL
NKTGEQVIVGHSQGTIGTIAFQIPELAKRIKMFALGPVAVAFCTSPNAKUGR
LPDLHLDKDFEFLPQSAFLWGLHVTCHVILKELCNLCLFLGCFERNLNMSR
VDVTTSPAGTSYQMLHMSQAVKFKQFQAFDGMSSAKNYFHYNOSYPPYTNKMDL
VPTAVSGSGHDLADVVDVNNILLTQITNLVFHESIPWEHLDFIWLGLDAPWRLNKKII
NLMEKYQ"
34. .96
sig_peptide
mat_peptide
variation
79
/product="lysosomal acid lipase"
/replacement="c"
100
variation
1234. .2480
3'UTR
polyA_signal
2459. .2464
polyA_site
2480
BASE COUNT 692 a 477 c 513 g 799 t
ORIGIN
Query Match 25.8%; Score 527.8; DB 9; Length 2481;
Best Local Similarity 66.6%; Pred. No. 2.7e-124;
Matches 751; Conservative 0; Mismatches 376; Indels 0; Gaps 0;
QY 240 AAAGCTGGACCCAGAGCAATCATGAATATTAGTGAATATCATCAACATCAAGGCTAT 299
Db 112 ACAGCTGGATCTCTGAAACAAACATGAATGTGAGTGAATATCTCTTCTGGGATTC 171
QY 300 CCCTGTGAGGATATGAGTGCACACTGAACATGGGTATATCTCTGTATACAGGATT 359
Db 172 CCTAGTGGAAATACCTAGTGTGAGACAGAGATGGATATATCTGTGCCCTTAACCGAAT 231
QY 360 CCTGAGGCTAGTGCAACCTAAGAAGACAGGTTCCAGGCTGTGGTGTACTGTCAGCAT 419
Db 232 CCTCATGGGAGGAGAACCACTTCTGACAAAGGTCCCAACCCAGTGTCTTCTCTGCAACAT 291
QY 420 GGCCTAGTGGAGGTCTAGCACTGAGATTTCCAACTGCCCAACATAGCTGGGCTTC 479
Db 292 GGCCTGCTGGCAGATCTAGTAACTGGGTCTCAAAACCTTGCACACAGCAGCTGGGCTTC 351
QY 480 ATTCTGCAGATGCTGGTGTGAGTGGGATGGGAAACAGCAGGGAAGCGCTGGTCT 539
Db 352 ATTCTGTGATGCTGGTGTGAGTGGGATGGGAAACAGCAGGGAATACCTGGTCT 411
QY 540 CGAAACACAGACACTCTCCATAGACCAAGATGAGTGTGGGCTTTTCAATATGATGAG 599
Db 412 CGGAAACATAAGACACTCTCAGTGTCTCAGGATGAATCTGGGCTTTTCAATATGATGAG 471
QY 600 ATGGCTAGGTTGACCTTCCTGAGTGAATATTTTTCAGAAACAGGCGCCAGAA 659
Db 472 ATGGCAAAATATGACCTACCACTTCCCAATTAATCTTCTGAATAAATCTGGCAGAA 531
QY 660 AAGATCTATTATGCGGTATTTCAGGCGCACCACTGGCTTTTATGCAATTTCCACC 719
Db 532 CAAGTGATATTATGCGGTATTTCAGGCGCACCACTATAGGTTTATAGCAATTTTCACAG 591
QY 720 ATGCCAGAGCTGGCTCAGAAATCAAAATGATTTTGTCTTTAGCACCATAGCCACGTT 779
Db 592 ATCCCTGAGCTGGCTTAAAGGATTAATGTTTGTCCCTGGGCTCTGGCTTCGCTC 651
QY 780 AAGCATGCAAAAGAGCCCGGACCAATTTTGTGTGCTGCCAGATATGATGATCAAGGCA 839
Db 652 GCCTTCTGCTAGCGCTTATGGCCAAATAGGACGATTAACGATCATCTCATTTAAGAC 711
QY 840 TTGTTTGGCAAAAGAAATTTCTGTATCAGACCAAGATTTCTCAGAACACTTGTATTAC 899
```

```
Db 712 TTATTGGAGACAAAGAAATTTCTCCCGAGTGGTGTGAGTGGCTGGTACCCAC 771
QY 900 CTTTGTGGCCAGGTGATTTCTTGATCAGATTTGTAGTAAATATCATGTTACTTCTGGGTGA 959
Db 772 GTTTCAGCTCATGTCTCATCTGAAGGAGCTCTGTGAAATCTCTGTGTTTCTTCTGTGTGA 831
QY 960 TTCAACACCAACAATATGAACATGATGAGCGAGCAAGTGTATATGTCGCCACACTCTTCT 1019
Db 832 TTTATGAGAGAAATTTAAATATGTTCTAGAGTGGATGTATATCAACACATCTTCTCTCT 891
QY 1020 GGAACATCTGTGCAAAATATTCTACCTGGAGCCAGGAGTGAATTTCTGGTGAATCTCGG 1079
Db 892 GGAACCTTCTGTGCAAAACATGTTTACCTGGAGCCAGGCTGTAAATTTCCAAAAGTTTCAA 951
QY 1080 GCATTGTGCTGGGAGTGAAGACCAAAATCTGGAATAATGCAATCAGCAACTCTCTCTA 1139
Db 952 GCCTTTGTGCTGGGAGCAGTGCAGGAATTTATTTTCTTACACCAAGTATCTCTCCC 1011
QY 1140 AGGTACAGAGTCAAGATATGACGGTCCCTACAGCAATGTGGACAGGAGTCAAGACTGG 1199
Db 1012 ACATACAATGTGAAGACATGCTTGTGCGACTGCAGTCTGGAGCGGGGTCACGACTGG 1071
QY 1200 CTTTCAATCCAGAGAGCTGAAATGCTGCTCTGTAGGTGACCAACTCATCTACCAT 1259
Db 1072 CTTGCAGATGTCTACGAGTCAATATCTTACTGACTCAGATCACCACCTTGGTGTTCAT 1131
QY 1260 AAGATATTTCTGAAATGGGCTCAGCTGAGTATTCATCTGGGGTTTGATGTCCTCACCCT 1319
Db 1132 GAGAGCATTTCCGAGTGGAGCATCTTCTGCTTCTTATTTGGGCCCTGGATGCCCTTGGAG 1191
QY 1320 ATGTCAATGAATCATCATCTGATGACAGGAGGAGAGCAACCT 1366
Db 1192 CTTTATAATAAATTTAATCTAATGAGGAATATCAGTGAAGCT 1238
```

RESULT 12

```
HSLAL
LOCUS HSLAL 2626 bp mRNA linear PRI 25-FEB-1994
DEFINITION H.sapiens mRNA for lysosomal acid lipase.
ACCESSION X76488
VERSION X76488.1 GI:434305
KEYWORDS lysosomal acid lipase.
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
REFERENCE 1
AUTHORS Ameis,D., Merkel,M., Eckerskorn,C. and Greten,H.
TITLE Purification, characterization and molecular cloning of human
JOURNAL hepatic lysosomal acid lipase
MEDLINE 9415897
PUBMED 8112342
REFERENCE 2 (bases 1 to 2626)
AUTHORS Ameis,D.
TITLE Direct Submission
JOURNAL Submitted (29-NOV-1993) D. Ameis, Medical Department, University
Hospital Eppendorf, Martinistrasse 52, 20246 Hamburg, FRG
FEATURES
source
1. .2626
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/chromosomes="10"
/tissue_type="liver"
1. .2626
/gene="LAL/LIPA"
146. .1345
/gene="LAL/LIPA"
/EC_numbers="3.1.1.13"
/notes="sterol esterase"
/codon_start=1
gene
CDS
```


/evidence=experimental	
/product="lysosomal acid lipase"	
/protein_id="CAA54026.1"	
/db_xref="GI:434306"	
/db_xref="SWISS-PROT:P38571"	
/translation="MKMRFLGVLCVILVTLHSEGGGKLTAVDPETNNVSEIISYV GPPSEYLVETEDGYTLCLNRIPIHGRKNSDKPKPVVFLQHLGLADSSNWNVLANS SLGFIILADAGFDVWGNSTWMSRKHKTLSVODEFWAFSYDENAKYDLPSINFLI NKGQBYVYVGHSSQTTIGTIPAFSQIPELAKRIKMFALGPVAVSACTSPMAKLR LPDLHKLDFGDKFIPQSAFLKWLIGVCTHVLKELCGNLFCFLCCGFERNLMSR VPTAVWSGGHDLADYIDVNIILLTOITNLVFHESIPWEHLDFIWLGLDAPRWLNKII NLMRKYQ"	
sig_peptide	146..226
/gene="LAL/LIPA"	
/evidence=experimental	
374..1342	
/gene="LAL/LIPA"	
/product="lysosomal acid lipase"	
/evidence=experimental	
227..373	
/gene="LAL/LIPA"	
/note="propeptide"	
/evidence=experimental	
BASE COUNT	741 a 527 c 547 g 811 t
ORIGIN	
Query Match 25.8%; Score 527.8; DB 9; Length 2626;	
Best Local Similarity 66.8%; Pred. No. 2.8e-124;	
Matches 751; Conservative 0; Mismatches 376; Indels 0; Gaps 0;	
Qy	240 AAGCTGTGACCCGAGAGCATTCATCAATATTAGTGAATCATCAACATCAAGCTAT 299
Db	224 ACAGCTGTGATCCTGAAACAAACATGAATGTGAGTGAATATCTCTACTGGGATTC 283
Qy	300 CCCTGTGAGAAATGAGTGCACCTGAGATGGGTATATCTTCTTTACAGGATT 359
Db	284 CCTAGTGAAGATACCTGATGTGAGACAGAGATGATATCTTGTGCTTACCGAAT 343
Qy	360 CCTCGAGGCTAGTGCACACCTAAGAGACAGGTTCAGGCTCTGTGTGTACTGACGAT 419
Db	344 CCTCATGGAGGAGAACCACTCTGACAAAGGTCCCAACACAGTGTCTCTCTGCAACAT 403
Qy	420 GGCTAGTGTGAGGTGTAGCACTGATTTCCAACTGCTCCCAACATGAGCTGGGCTTC 479
Db	404 GGCTGTGTCAGATCTTAGTAACTGGGTACAAACCTTGGCCACAGAGCTGGGCTTC 463
Qy	480 ATTCTGCAGATCTGTGTTTGTGATGGTGGGAAACAGAGGGGAAACGCTGGTCT 539
Db	464 ATTCTGTGATCTGTGTTTGTGATGGTGGGAAACAGAGGGGAAATACCTGGTCT 523
Qy	540 CGAAACACAGACACTCTCCATAGACCAAGATGATTTCTGGCTTTTCAATGATGAG 599
Db	524 CGGAAACATAAGACACTCTCAGTTCTCAGGATGAATTTCTGGCTTTTCAATGATGAG 583
Qy	600 ATGCTAGTGTGACCTCTGCTGATGATGAATTTTCTGAGAAACGGGCGGAGAA 659
Db	584 ATGCAAAATATGACCTACCACTGCTTCAAGGACCACTATAGTTTATAGCAATTTTCACAG 643
Qy	660 AAGATCTATTATGTGCGCTATTACAGGGCACCACCATGGCTTTTATGCAATTTTCCAC 719
Db	644 CAAGTGATTTATGTGGGTCTATTCTCAAGGACCACTATAGTTTATAGCAATTTTTCACAG 703
Qy	720 ATGCGAGAGTGGCTCAGAAATCAAAATGATTTTCTGTCGACATGATGATCAAGGGA 779
Db	704 ATCCCTGAGTGGCTTAAAGGATTAAGATTTTCTGCTGGCTCTGCTGGCTTCGCTC 763
Qy	780 AAGCATGCAAAAGCCCGGACCAATTTTCTGTCGACATGATGATCAAGGGA 839
Db	764 GCCTTCTGTACTAGCCCTATGGCCAAATATTAGGACGATACCAATCATCAATAGGAC 823
Qy	840 TTGTTGGCAAAAGAAATTTCTGTATCAGACAGATTTCTCAGACAACTGTGATTTAC 899

Db	824 TTATTTGGACACAAAGAAATTTCTTCCCGAGAGTGGCTTTTGAAGTGGTGGTACCCAC 883
Qy	900 CTTTGTGGCCAGGTGATTTCTGATCAGATTTGTAGTAAATATCATGTATCTTCTGGGTGA 959
Db	884 GTTTGCACTCATGTCTACTGAAGAGGAGTCTGTGGAATATCTCTGTTTCTTCTGTGTGA 943
Qy	960 TTCAACACCAACATATGAACATGAGCCGAGCAAGTGTATATGCTGCCACACACTCTTGCT 1019
Db	944 TTTAATGAGAGAAATTTAATATATGCTAGAGTGGATATATACACACATCTCTCTGCT 1003
Qy	1020 GGAACATCTGTGCAAAATATTTACACTGGAGCAGGAGTGAATTTCTGGTGAACCTCGG 1079
Db	1004 GGAACCTCTGTGCAAAACATGTTACACTGGAGCCAGGCTGTTAAATTCAAAAGTTTCAA 1063
Qy	1080 GCATTTGACCTGGGAGTGGAGACCAAAATCTGAAAATATGCAATCAGCCACTCTCTGTA 1139
Db	1064 GCCTTTGACTGGGAGCAGTGCAGAGATATTTTTCATTACACAGAGTATCTCTCC 1123
Qy	1140 AGGTACAGAGTTCAGAGATATGACGGTCCCTACAGCAATGTGGACAGAGGTCAAGACTGG 1199
Db	1124 ACATACAAATGTGAAGGACATGCTTGTGCCGACTGCTGAGCGGGGTCAAGACTGG 1183
Qy	1200 CTTTCAATCCAGAGACGTGAAATGCTCTCTCTGAGTGACCAACCTCATCTACCAT 1259
Db	1184 CTTGCAATGTTACGACGTCAATATCTTACTGACTCAGATCACCACCTTGTGTCCAT 1243
Qy	1260 AAGAAATATCTGAATGGGCTCAGTGGATTTTCATCTGGGGTTTGGATGCTCCTCACCGT 1319
Db	1244 GAGAGCATTCGGAATGGGAGCATTTGACTTCTATTGGGGCTGGATGCCCTTGGAGG 1303
Qy	1320 ATGTACAAATGAATCATCTCATCTGATGACAGAGGAGGACCAACCT 1366
Db	1304 CTTTATAATAAATTAATTAATCTAATGAGGAAATATCAGTGAAGCT 1350
RESULT 13	
BD094072	
LOCUS	
DEFINITION	
ACCESSION	
VERSION	
KEYWORDS	
SOURCE	
ORGANISM	
REFERENCE	
AUTHORS	
TITLE	
JOURNAL	
COMMENT	
1 (bases 1 to 2493) Nojima, H., Yoshisue, H., Obayashi, M., Ota, T., Kawabata, A., Sakurada, K., Kuga, T., Sekine, S., Nakamura, Y. and Sugan, S. Shear stress-responsive DNAs Patent: WO 0125427-A 33 12-APR-2001; KYOWA HAKKO KOGYO CO LTD, HIROSHI NOJIMA, HAJIME YOSHISUE, MASAYA OBAYASHI, TOSHIO OTA, AYAKO KAWABATA, KAZUHIRO SAKURADA, TETSURO KUGA, SUSUMU SEKINE, YUSUKE NAKAMURA, SUMIO SUGANO OS Homo sapiens (human) PN WO 0125427-A/33 PD 12-APR-2001 PF 02-OCT-2000 WO 2000JP006840 PR 01-OCT-1999 JP 99P 280976 PI HIROSHI NOJIMA, HAJIME YOSHISUE, MASAYA OBAYASHI, TOSHIO OTA, PI AVAKO KAWABATA, PI KAZUHIRO SAKURADA, TETSURO KUGA, SUSUMU SEKINE, YUSUKE NAKAMURA, PI SUMIO SUGANO PC C12H15/12,C07K14/435,C07K16/18,C12P21/02,C12Q1/68,A61K38/00, PC A61K39/395, PC A61K48/00,A61P9/10,G01N33/50,G01N33/53 CC Key Location/Qualifiers CDS (41)..(1237). 1..2493 /organism="Homo sapiens" /mol_type="genomic DNA"	
FEATURES	
source	


```

/genes="LIPA"
/notes="G00-120-153"
2493
polya_site
/genes="LIPA"
/notes="G00-120-153"
692 a 484 c 516 g 801 t
BASE COUNT
ORIGIN
Query Match 25.7%; Score 526.2; DB 9; Length 2493;
Best Local Similarity 66.6%; Pred. No. 7.1e-124;
Matches 751; Conservative 0; Mismatches 376; Indels 0; Gaps 0;
Qy 240 AAGCTGTGACCCAGCAAGCATTCATGAATATTAGTGAATCATCAACATCAAGCTAT 299
Db 119 ACAGCTGTGATCCTGAAACAAACATCAATGTGAGTGAATATCTCTTACTGGGATTC 178
Qy 300 CCCTGTGAGGAATATGAAGTCGCAACTGAAGTGGGTATATCTTCTGTACAGGATT 359
Db 179 CTAAGTGAGGAATACCTAGTTGAGACAGAGATGGATATATCTGTGCTTACCGAATT 238
Qy 360 CCTCGAGGCTAGTCAACCTAAGAAAGACAGAGTTCAGGCTGTGTGTGTACTGACGAT 419
Db 239 CCTCATGGGAGGAAGAACCATCTCTGACAAAGGTCCCAACACAGTTCTCTCTGCAACAT 298
Qy 420 GGCTAGTTGGAGTGTAGCACTGATTTCCAACTGCTGCCCAACATAGCTGGCTTC 479
Db 299 GGCTTGTGCGAATTTCTAGTAATCTGGTCCAAACCTTGCACACAGCAGCTGGCTTC 358
Qy 480 ATTCTGCGACATGTGTTTGTGCTGTGATGGGAAACAGAGGGGAAACGCTGGTCT 539
Db 359 ATCTTGTCTGATGCTGTTTGTGCTGTGATGGGAAACAGAGGAAATACCTGGTCT 418
Qy 540 CGAAACACAGACACCTCTCCATAGACCAAGATGAGTTCTGGCTTTCATGTATGATGAG 599
Db 419 CGAAACACATAGACACCTCTCAGTTCTCAGAGTGAATTTCTGGCTTTCATGTATGATGAG 478
Qy 600 ATGCTAGTTGACCTCTCTGAGTGAATTAATTTTTCAGAAACAGGAGGAGGAGGAGGAA 659
Db 479 ATGCAAAATATGACCTACAGCTTCCATTAATTTCTGATTAATTAATTAATTAATTAAT 538
Qy 660 AAGATCTATTATGCTGCTATTCAGAGGACCAACAGGCTTTTATGCAATTTTCCACC 719
Db 539 CAAGTGTATTATGCTGCTATTCAGAGGACCAACAGGCTTTTATGCAATTTTCCACC 598
Qy 720 ATGCGAGCTGCTGCAAAATCAAAATGATTTTGTGCTGCGAGATATGATCAAGGGA 779
Db 599 ATCCCTGAGCTGGCTAAAGATTAATAATGTTTGTGCTGCGAGATATGATCAAGGGA 658
Qy 780 AAGCATGCAAAAGCCCGGACCAAAATTTTGTGCTGCGAGATATGATCAAGGGA 839
Db 659 GCCTTCTGTACTAGCCCTATGGCCAAATTAGGAGGATTAACAGATCATCTCATTAAGGAC 718
Qy 840 TTGTTGGCAAAAGAAATTTCTGTATCAGACCGAGTTTCTCAGACAACTGTGTATTATAC 899
Db 719 TTATTTGGAGCAAAAGAAATTTCTTCCACAGATGCGTGTGTTTGAAGTGGCTGAGTACCCAC 778
Qy 900 CTTTGTGCGCAGGATTTCTTGATCAGATTTGTAGTAATATCATGTTTACTTCTGGGTGGA 959
Db 779 GTTTGCACTATGCTACTGAGGAGCTCTGTGGAATCTCTGTTTCTTCTGTGTGGA 838
Qy 960 TTCACACCAACAAATATGAACATAGCGGAGCAAGTGTATATGCTGCGCCACATCTTGTCT 1019
Db 839 TTTAATGAGAGAAATTTAAATATGCTAGAGTGGATGTATATACACACATTTCTCTGTCT 898
Qy 1020 GGAACNTGTGCAAAATTTCTACATGCGGAGCCAGCAGTGAATCTGCTGTAAGTCTCCGG 1079
Db 899 GGAATTTCTGTGCAAAATTTCTACATGCGGAGCCAGCAGTGTAAATTTCCAAAGTTTCAA 958
Qy 1080 GCATTTGACTGGGAGTGAACCAAAATCTGAAATAATGCAATCAGCACTCTCTGTA 1139
Db 959 GCCTTTGACTGGGAGTGAACCAAAATTTCTGAAATAATGCAATCAGCACTCTCTGTA 1018
Qy 1140 AGGTACAGAGTCAGAGATATGACGGTCCCTACAGCAATGTGGACAGAGGTGAGGACTGG 1199

```

1019 ACATCAATGTGAAGGACATGCTTGTGCCACTGCACTGTGGAGCGGGGTCAAGACTGG 1078

1200 CTTTCAATCCAGAGAGCGTGAATGCTCTCTCTGAGTGACCAACCTCATCTACCAT 1259

1079 CTTGAGATGCTACGAGGCAATATCTTACTGACTCAGATCAGATCAGGCTGGTTCAT 1138

1260 AAGAATATTCCTGAATGGGCTCAGTGGATTTTCATCTGGGGTTTGATGCTCTCTACCGT 1319

1139 GAGAGCATTCGGAATGGGAGCATCTTGAATTCATTTGGGGCTGGATGCCCTTGGAGG 1198

1320 ATGTACATGAATCATCTGATGATGACGAGGAGGAGACCACT 1366

1199 CTTTATAATAAATTTATTAATCTAATGAGGAATATCAGTGAAGCT 1245

RESULT 15

BC012287 2586 bp mRNA linear PRI 08-AUG-2001

LOCUS Homo sapiens, similar to lipase A, lysosomal acid, cholesterol

DEFINITION esterase (Wolman disease), clone MGC:5229 IMAGE:2900168, mRNA,

complete cds.

ACCESSION BC012287 GI:15126726

VERSION BC012287

KEYWORDS MGC.

SOURCE Homo sapiens (human)

ORGANISM Homo sapiens

REFERENCE 1 (bases 1 to 2586)

AUTHORS Strausberg, R.

TITLE Direct Submission

JOURNAL Submitted (08-AUG-2001) National Institutes of Health, Mammalian

Gene Collection (MGC), Cancer Genomics Office, National Cancer

Institute, 31 Center Drive, Room 11A03, Bethesda, MD 20892-2590,

USA

REMARK NIH-MGC Project URL: <http://mgc.nci.nih.gov>

COMMENT Contact: MGC help desk

Email: cgapbs-remail.nih.gov

Tissue Procurement: ATCC

cDNA Library Preparation: Life Technologies, Inc.

cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)

DNA Sequencing by: Baylor College of Medicine Human Genome

Sequencing Center

Center code: BCM-HGSC

Web site: <http://www.hgsc.bcm.tmc.edu/cdna/>

Contact: villalob@bcm.tmc.edu

Villalob, D.K., Luna, R.A., Hale, S.M., Hulyk, S., Lu, X., Garcia,

A.M., Holloway, M., Telford, B., Hodgson, A., Bouck, J., Yu, W.,

Muzny, D.M., Gibbs, R.A.

Clone distribution: MGC clone distribution information can be found

through the I.M.A.G.E. Consortium/LLNL at: <http://image.llnl.gov>

Series: IRAP Plate: 3 Row: c Column: 3

This clone was selected for full length sequencing because it

passed the following selection criteria: matched mRNA gi: 434305.

Location/Qualifiers

1. 2586

/organism="Homo sapiens"

/mol_type="mRNA"

/db_xref="taxon:9606"

/clone="MGC:5229 IMAGE:2900168"

/tissue_type="Placenta, choriocarcinoma"

/clone_lib="NIH_MGC_10"

/lab_host="DH10B"

/note="vector: pCMV-SPORT6"

101..1300

/codon_start=1

/product="Similar to lipase A, lysosomal acid, cholesterol

esterase (Wolman disease)"

/protein_id="AAH12287.1"

/db_xref="GI:15126726"

/translation="MKRFLGLVYVCLVLPVHSEGGKLTALDPETNNVSEIISVM"

FEATURES

source

CDS

GFPSSEYLVETEDGVILCLNRI PHGRKNHSDKPKVPVFLQHLGLADSNVNTLANS
SLGRLADAGDVMSRSGNTWRSKKTLSVDSODEFAPSWYEMAKYDLPASINPL
NKTQBOVYVYVHSGQTTIGFIAFSQJPELAKRIKMFALGPVASFCTSPAKLGR
LPDLHLDKDFKDFLQSALFKWLFHVCVHVLKELCNELCFLLCGFRNERNLSR
VDTTTSFAGTSVQNMHLWSQAVKQFQAFDMGSSAKNYFHYNQSYPTYNVKML
VPTAVMSGGHMDLADVVDVNIILLTOINLVPHESIPEWEHLDFWGLDAPWRLYNKI
NLMRYQ

BASE COUNT 736 a 511 c 531 g 808 t

Query Match 25.7%; Score 524.6; DB 9; Length 2586;
Best Local Similarity 86.5%; Pred. No. 1.8e-127;
Matches 750; Conservative 0; Mismatches 327; Indels 0; Gaps 0;

QY	240	AAAGCTGTGGACCCAGAGCAATTCATGAATATTAGTGAATCATCAACATCAAGGCTAT	299
DB	179	ACAGCTTTGGATCCTGAACAACAACATGAATGTGAGTGAATATCTCTTACTGGGGATTC	238
QY	300	CCCTGTGAGGAATATGAAGTGCBACTGAAGATGGGTATATCTTTCTGTGTAAAGGATT	359
DB	239	CCTAGTGAGGAATACCTAGTTGAGACAGAGATGGATATATTCTGTGCTTAACCGAATT	298
QY	360	CCTGAGGCTTAGTGAACCTAAGAGACAGGTTCCAGGCTGTGGTGTACTGCGCAT	419
DB	299	CCTCATGGAGGAGAACCATCTGCAAAAGTCCAAACAGTTGCTTCTCTGCAACAT	358
QY	420	GGCCTAGTTGAGGTGTAGCAACTGGATTTCACACCTGCCCAACAATAGCCTGGGCTTC	479
DB	359	GGCTGTGCGCAGATTCTAGTAACTGGGTCAACAACCTTGCCACAGCAGCCTGGGCTTC	418
QY	480	ATTCCTGGCAGATGCTGGTTTGAAGTGTGGATGGGACAGCAGGGGAAACGCTGGTCT	539
DB	419	ATTCCTGTGATGCTGGTTTGAAGTGTGGATGGGACAGCAGGGGAAATACCTGGTCT	478
QY	540	CGAAACCAACAGACACTCTCCATAGACCAAGATGAGTTCTGGGCTTTTCAGTTATGATGAG	599
DB	479	CGGAACATAAGACACTCTCAGTTTCTCAGGATGAATTCGGGCTTTCAGTTATGATGAG	538
QY	600	ATGGCTAGGTTGACCTTCCTGCGAGTGATAACTTTATTTTGCAGAAACCGGGCCAGAA	659
DB	539	ATGGCAAAATATGACCTACCACTGCTCCATTAACCTTCAATCTGAATAAACTGGCCAAAG	598
QY	660	AAGATCTATTATGCTGGCTATTACAGGGCCACCACTGGCTTTTATGCAATTTTCCACC	719
DB	599	CAAGTGATATATGCTGGCTATTCTCAAGCACCCTATAGTATTTTATAGCAATTTTCAG	658
QY	720	ATGCCAGAGCTGGCTCAGAAATCAAAATGATTTTGTCTTTAGCACCCTATAGCCACTGTT	779
DB	659	ATCCCTGAGCTGGCTTAAAGGATTAATGTTTGTGGCTGGCTCTGCTGGCTTCGCTC	718
QY	780	AAGCATGCAAAAGCCCGGACCAAAATTTTGTGCTGCCAGATATGATGATCAAGGA	839
DB	719	GCCTTCTGTACTAGCCCTATGGCCAAATTAGGACGATTACCAGATCATCTCAATTAGGAC	778
QY	840	TTGTTTGGCAAAAGAAATTTCTGATCAGACACAGATTTCTCAGACAACTGTTTATTATAC	899
DB	779	TTATTTGGAGACAAGAAATTTCTCCCGAGAGTGGTTTGAAGTGGCTGGGTACCCAC	838
QY	900	CTTTGTGGCCAGGTGAATCTTGATCAGATTTGTAGTAATATCATGTTACTTCTGGGTGA	959
DB	839	GTTTGCACTCATGTACTAGGAGGCTCTGTGGAATCTCTGTTTCTTCTCTGTGTGA	898
QY	960	TTCAACACCAACAATATGACATGAGCGGAGTGTATATGCTGCCACACTCTTGTCT	1019
DB	899	TTTATGAGAGAAATTTAAATATGCTAGAGTGGATGATATACAAACATCTCTCTGCT	958
QY	1020	GGAACTCTGTGCAAAATATTCTACCTGGAGCCAGGAGTGAATTTCTGTGAACCTCCG	1079
DB	959	GGAACTCTGTGCAAAACATGTTTACCTGGAGCCAGGCTGTTAAATCCAAAAGTTTCAA	1018
QY	1080	GCATTTGACTGGGAGTGAGACCAAAATCTGGAAAATGCAATCAGCCAACTCTCTGTA	1139
DB	1019	GCCTTTGACTGGGGAAGCAGTGCCCAAGAAATTTTTCATTACCAACAGATTATCTCTCC	1078

Search completed: February 19, 2004, 08:29:14
Job time : 7525.62 secs

QY	1140	AGTACAGAGTCAGATATGACGGTCCCTACAGCAATCTGGACAGGAGGTCAGACTGG	1199
DB	1079	ACATACATGTGAAGGACATGCTTGTGCCGACTGCAGTCTGGAGCGGGGTCAAGACTGG	1138
QY	1200	CTTTCAAATCCAGAAAGACGTGAAAATGTGCTCTCTGAGGTGACCAACCTCATCTTACCAT	1259
DB	1139	CTTGCAGATGTCTAGCAGCTCAATATCTTACTGACTCAGATCACCACCTTGGTGTCCAT	1198
QY	1260	AAGATATTCTGATGGGCTCACTGGATTTTCATCTGGGTTTGGATGCTCTCACCGT	1319
DB	1199	GAGAGCATTTCCGGAATGGAGCATCTTGACTTTCAATTTGGGGCTGGATGCCCTTTGGAGG	1258
QY	1320	ATGTACAATGAATATCATCTCATCTGATGACAGCAGGAGGAGACCAACCT	1366
DB	1259	CTTTATAATAAATTTATTAATCTTAATGAGGAATATCAGTGAAGCT	1305

GenCore version 5.1.6
Copyright (c) 1993 - 2004 CompuGen Ltd.

OM nucleic - nucleic search, using sw model

Run on: February 19, 2004, 04:41:17 ; Search time 563.905 Seconds
(without alignments)
9784.725 Million cell updates/sec

Title: US-10-042-431-45
Perfect score: 2044
Sequence: 1 gtcgaccacgctgcgggg.....aaaaaaagggcgccgc 2044

Scoring table: IDENTITY NUC
Gapop 10.0 , Gapext 1.0

Searched: 2552756 seqs, 1349719017 residues

Total number of hits satisfying chosen parameters: 5105512

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : N_Geneseq_19Jun03;+
1: /SIDSI/gcgdata/geneseq/geneseq-emb1/NA1980.DAT.*
2: /SIDSI/gcgdata/geneseq/geneseq-emb1/NA1981.DAT.*
3: /SIDSI/gcgdata/geneseq/geneseq-emb1/NA1982.DAT.*
4: /SIDSI/gcgdata/geneseq/geneseq-emb1/NA1983.DAT.*
5: /SIDSI/gcgdata/geneseq/geneseq-emb1/NA1984.DAT.*
6: /SIDSI/gcgdata/geneseq/geneseq-emb1/NA1985.DAT.*
7: /SIDSI/gcgdata/geneseq/geneseq-emb1/NA1986.DAT.*
8: /SIDSI/gcgdata/geneseq/geneseq-emb1/NA1987.DAT.*
9: /SIDSI/gcgdata/geneseq/geneseq-emb1/NA1988.DAT.*
10: /SIDSI/gcgdata/geneseq/geneseq-emb1/NA1989.DAT.*
11: /SIDSI/gcgdata/geneseq/geneseq-emb1/NA1990.DAT.*
12: /SIDSI/gcgdata/geneseq/geneseq-emb1/NA1991.DAT.*
13: /SIDSI/gcgdata/geneseq/geneseq-emb1/NA1992.DAT.*
14: /SIDSI/gcgdata/geneseq/geneseq-emb1/NA1993.DAT.*
15: /SIDSI/gcgdata/geneseq/geneseq-emb1/NA1994.DAT.*
16: /SIDSI/gcgdata/geneseq/geneseq-emb1/NA1995.DAT.*
17: /SIDSI/gcgdata/geneseq/geneseq-emb1/NA1996.DAT.*
18: /SIDSI/gcgdata/geneseq/geneseq-emb1/NA1997.DAT.*
19: /SIDSI/gcgdata/geneseq/geneseq-emb1/NA1998.DAT.*
20: /SIDSI/gcgdata/geneseq/geneseq-emb1/NA1999.DAT.*
21: /SIDSI/gcgdata/geneseq/geneseq-emb1/NA2000.DAT.*
22: /SIDSI/gcgdata/geneseq/geneseq-emb1/NA2001A.DAT.*
23: /SIDSI/gcgdata/geneseq/geneseq-emb1/NA2001B.DAT.*
24: /SIDSI/gcgdata/geneseq/geneseq-emb1/NA2002.DAT.*
25: /SIDSI/gcgdata/geneseq/geneseq-emb1/NA2003.DAT.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	2044	100.0	2044	22 AAF45131	Human TANGO 294 cD
2	2044	100.0	2044	25 ABK94106	cDNA encoding huma
3	1993.8	97.5	2035	22 AAD19226	Human CG162 (or C5
4	1343.4	65.7	1384	22 AAD19226	Human CG162 (or C5
5	1269	62.1	1269	22 AAF45132	Human TANGO 294 OR
6	1138	55.7	1210	24 ABK96569	Human cDNA encodin
7	1126	55.1	1206	24 ABS55356	cDNA encoding huma
8	1107.8	54.2	1272	24 ABK90871	cDNA encoding huma

9	890.4	43.6	18554	24 ABS55357	Genomic DNA encodi
10	752	36.8	1116	24 ABK96570	Human cDNA encodin
11	527.8	25.8	2481	24 ABK97363	Gene #3881 used to
12	526.2	25.7	2493	22 AAH02906	Human shear stress
13	489.2	23.9	1378	14 AAQ42310	RGL precursor. Or
14	470	23.0	1367	7 AAN60685	Sequence encoding
15	470	23.0	1367	7 AAN60566	Sequence encoding
16	470	23.0	1367	17 AAT58916	Human gastric lipa
17	468.4	22.9	1695	22 AAH57420	Human stomach cell
18	464.4	22.7	1336	6 AAN50385	Rat prelingual lip
19	456.8	22.3	1411	24 ABL57144	Kid goat pregastric
20	454.2	22.2	1194	24 AAD27803	Human lysosomal ac
21	454	22.2	1228	24 ABK12385	cDNA encoding huma
22	450.8	22.1	1284	24 ABK85978	Yeast YB-1 express
23	449.8	22.0	8324	24 ABL57145	Human encoding huma
24	448.4	21.9	1473	24 ABK12388	cDNA encoding huma
25	446.8	21.9	1185	24 AAD31193	Human triacylglyce
26	444.2	21.7	1528	15 AAQ68388	Canine gastric lip
27	444.2	21.7	1528	17 AAT58915	Dog gastric lipase
28	444.2	21.7	1531	15 AAQ68389	Canine gastric lip
29	443.6	21.7	1200	22 AAH78205	Nucleotide sequenc
30	424.8	20.8	1260	24 ABK1703	Human encoding huma
31	419.8	20.5	1086	25 ACC44783	Human triacylglyce
32	414.2	20.3	1360	24 AAL47480	Human lysosomal ac
33	414.2	20.3	1360	25 ABK14880	Human lipase cDNA.
34	374.4	18.3	996	25 ACC46563	Human dithp secret
35	346.4	16.9	1138	24 ABQ82326	Human NOV1 encodin
36	332.8	16.3	1098	24 AAD27802	Human lysosomal ac
37	319.4	15.6	1143	24 ABK12386	cDNA encoding huma
38	315.4	15.4	828	24 AAD31194	Human triacylglyce
39	300.4	14.7	1377	24 AAD31192	Human encoding huma
40	290	14.2	842	24 AAD31192	Human triacylglyce
41	286	14.0	699	22 AAF28679	Human protein HP03
42	286	14.0	1308	22 AAF28689	Human protein HP03
43	282.8	13.8	802	22 AAI97406	Human neuroblastom
44	257.8	12.6	617	21 AAA44349	Human secreted exp
45	244.6	12.0	1971	23 AAS91026	DNA encoding novel

ALIGNMENTS

RESULT 1

AAF45131
ID AAF45131 standard; cDNA; 2044 BP.

XX AAF45131;

XX AAF45131;

DT 30-MAR-2001 (first entry)

XX Human TANGO 294 cDNA.

XX Gene therapy; TANGO protein; INTERCEPT protein; neurological disorder;
XX central nervous system; focal brain disorder; bipolar affective disorder;
XX global-diffuse cerebral disorder; cerebrovascular; Alzheimer's disease;
XX senile dementia; Huntington's disease; amyotrophic lateral sclerosis;
XX Parkinson's; Gilles de la Tourette's syndrome; hypertension; sleep;
XX neurophysiologic; psychoactive substance use; anxiety; ss.
XX Homo sapiens.

XX WO200077239-A2

XX 21-DEC-2000

XX 24-MAY-2000; 2000WO-US14858.

XX 14-JUN-1999; 99US-0333159

XX (MILITARY) MILLENNIUM PHARM INC.

XX McCarthy SA, Fraser CC, Sharp JD, Barnes TM;

XX

Db 1741 AAATTGGGACAGATATTGAGTCTGGAGTCTGGATTATTTGTTGACATAAATTA 1800
 Qy 1801 GCTAGACATTTTACCTTGTGTCACAGAGACATACCTACCTCAGGAGCTGAGTGC 1860
 Db 1801 GCTAGACATTTTACCTTGTGTCACAGAGACATACCTACCTCAGGAGCTGAGTGC 1860
 Qy 1861 TTTAAGGACACACAAACAAATCAGTGTACAGTATGAGTAAATCTATGTTAAGCAIT 1920
 Db 1861 TTTAAGGACACACAAACAAATCAGTGTACAGTATGAGTAAATCTATGTTAAGCAIT 1920
 Qy 1921 CTCAGATTAAGCCCAAGTTTATAGTTGCACTCTCAGGAGAGAAAATTTATAGATGTTT 1980
 Db 1921 CTCAGATTAAGCCCAAGTTTATAGTTGCACTCTCAGGAGAGAAAATTTATAGATGTTT 1980
 Qy 1981 ATGAGTTCTCCAAATAATGCAATCTGCATTAACATTAATAAAAAAAAAAAAAAGGCGG 2040
 Db 1981 ATGAGTTCTCCAAATAATGCAATCTGCATTAACATTAATAAAAAAAAAAAAAAGGCGG 2040
 Qy 2041 CCGC 2044
 Db 2041 CCGC 2044

RESULT 2

ABX94106
 ID ABX94106 standard; cDNA; 2044 BP.
 AC ABX94106;
 XX
 DT 03-JUN-2003 (first entry)
 XX
 DE cDNA encoding human TANGO 294.
 XX

Human; TANGO 294; INTERCEPT; cellular process; tissue typing;
 forensic biology; cellular protease activity; cell interaction;
 development; blood disorder; haematopoietic cell-related disorder;
 growth; cell proliferation; cell differentiation; gamma delta T-cell;
 immune system; metabolic disorder; homeostatic disorder; anaemia;
 developmental bone disorder; osteoporosis; bacterial infection;
 thrombocytopenia; renal failure; liver disease; cystic fibrosis;
 obesity; cancer; bronchitis; asthma; emphysema; pulmonary oedema;
 respiratory distress syndrome; osteopathic; antibacterial; antianemic;
 thrombolytic; nephrotropic; antibiotic; hepatotropic; cytostatic;
 antiinflammatory; antiasthmatic; gene; ss.
 XX Homo sapiens.
 XX
 Key Location/Qualifiers
 FH 126..1397
 CDS /*tag= a
 FT /product= "TANGO 294"
 FT /note= "The ORF given as SEQ ID No:46 is specifically
 FT claimed in Claim 2"
 XX
 PN US2002182675-A1.
 XX
 PD 05-DEC-2002.
 XX
 PF 25-OCT-2001; 2001US-0042431.
 XX
 PR 14-JUN-1999; 99US-0333159.
 PR 24-MAY-2000; 2000US-0578063.
 XX
 PA (MILL-) MILLENNIUM PHARM INC.
 XX
 PI McCarthy SA, Fraser CC, Sharp JD, Barnes TM;
 XX
 DR WPI; 2003-328617/31.
 DR P-PSDB; ABU08369.
 XX
 PT New TANGO and INTERCEPT proteins, useful as modulating agents in
 regulating a variety of cellular processes, in chromosome mapping, in

PT tissue typing, and in forensic biology -
 XX Claim 2; Fig 6A-6C; 232pp; English.
 XX
 CC The present invention relates to the isolation of novel TANGO or
 CC INTERCEPT proteins, and the polynucleotide sequences encoding them.
 CC The invention discloses sequences for human TANGO 202, TANGO 234,
 CC TANGO 265, TANGO 273, TANGO 286, TANGO 294 and INTERCEPT 296. The
 CC sequences for murine TANGO 202 and TANGO 273 are also provided. The
 CC TANGO polypeptides and the polynucleotide sequences encoding them are
 CC useful as modulating agents in regulating a variety of cellular
 CC processes. The polynucleotide sequences are useful as primers or
 CC hybridisation probes for the detection of nucleic acids encoding
 CC TANGO polypeptides, chromosome mapping, tissue typing, and in
 CC forensic biology. TANGO 202 can be used to diagnose, prevent or
 CC treat disorders relating to aberrant cellular protease activity,
 CC development, and blood and haematopoietic cell-related disorders.
 CC TANGO 234 can be used to modulate growth, proliferation, survival,
 CC differentiation, and activity of gamma delta T-cells. TANGO 265 can
 CC be used to prevent, diagnose and treat disorders characterised by
 CC aberrant organisation or development of a tissue or organ, and for
 CC modulating differentiation of cells of the immune system. TANGO 273
 CC is useful for diagnosing, treating or preventing e.g. metabolic,
 CC homeostatic and developmental bone disorders (e.g. osteoporosis),
 CC and bacterial infection. TANGO 286 is useful for treating or
 CC preventing e.g. anaemia, thrombocytopenia, renal failure or
 CC liver disease. TANGO 294 is useful for treating or preventing
 CC e.g. cystic fibrosis or obesity. INTERCEPT 296 is useful for
 CC diagnosing, treating or preventing e.g. cancers, bronchitis, cystic
 CC fibrosis, asthma, emphysema, pulmonary oedema, or adult and infant
 CC respiratory distress syndromes. The present sequence encodes
 CC human TANGO 294.
 XX

Sequence 2044 BP; 618 A; 401 C; 460 G; 565 T; 0 other;
 Query Match 100.0%; Score 2044; DB 25; Length 2044;
 Best Local Similarity 100.0%; Pred. No. 0;
 Matches 2044; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 Qy 1 GTCGACCCACGCGTCCGGGAAATTCACAGAGAAATATGTAAGAGTTTTAAACCCAC 60
 Db 1 GTCGACCCACGCGTCCGGGAAATTCACAGAGAAATATGTAAGAGTTTTAAACCCAC 60
 Qy 61 AAATTTCTTCTTACTTTAGAAATTTAGTTTACATTCGCGAGGAAATAAATGCGATGTT 120
 Db 61 AAATTTCTTCTTACTTTAGAAATTTAGTTTACATTCGCGAGGAAATAAATGCGATGTT 120
 Qy 121 GGACCATGTTGGAAACCTTGTCAAGACAGTGGATTGTCTCACAGAAATGGAATGTGGC 180
 Db 121 GGACCATGTTGGAAACCTTGTCAAGACAGTGGATTGTCTCACAGAAATGGAATGTGGC 180
 Qy 181 TTCTGATTTCTGGTGGCGTATATGTTCCAGAGAAATGTAATTCAGTACATATGCCAATA 240
 Db 181 TTCTGATTTCTGGTGGCGTATATGTTCCAGAGAAATGTAATTCAGTACATATGCCAATA 240
 Qy 241 AAGCTGTGGACCCAGAGCAATTCATGAATATATAGTGAATCATCCACATCAAGCTATC 300
 Db 241 AAGCTGTGGACCCAGAGCAATTCATGAATATATAGTGAATCATCCACATCAAGCTATC 300
 Qy 301 CCTGTGAGGAATATGAAGTCCGAATCGAAGTGGTATATCTTCTTCTTACAGGATTC 360
 Db 301 CCTGTGAGGAATATGAAGTCCGAATCGAAGTGGTATATCTTCTTCTTACAGGATTC 360
 Qy 361 CTCGAGGCTTAGTGCAACCTTAAGAGACAGGTTCCAGGCTGTGTGTACTTCGAGCATG 420
 Db 361 CTCGAGGCTTAGTGCAACCTTAAGAGACAGGTTCCAGGCTGTGTGTACTTCGAGCATG 420
 Qy 421 GCTAGTTGGAGGTGCTAGCAACTGGAATTTCCAACTGCCCAACAATAGCCTGGGCTTCA 480
 Db 421 GCTAGTTGGAGGTGCTAGCAACTGGAATTTCCAACTGCCCAACAATAGCCTGGGCTTCA 480
 Qy 481 TTCTGCGCATGCTGGTTTTTGACGTTGGATGGGGAACAGCAGGGGAAACGCTGTCTC 540

Db 481 TTTCTGGCAGATGCTGGTTTGAAGTGGATGGGAAACAGACAGCGGAAACCGCTGCTC 540
 QY 541 GAAACACAGACACTCTCCATAGACCAAGATGAGTTCTGGGCTTTTCAGTTATGATGAGA 600
 Db 541 GAAACACAGACACTCTCCATAGACCAAGATGAGTTCTGGGCTTTTCAGTTATGATGAGA 600
 QY 601 TGGCTAGGTTTGAACCTTCTGCAAGTATGATAAATTTTATTTTGCAGAAAACGGGCGAGGAAA 660
 Db 601 TGGCTAGGTTTGAACCTTCTGCAAGTATGATAAATTTTATTTTGCAGAAAACGGGCGAGGAAA 660
 QY 661 AGATCTATATGTCGGCTATTCACAGGGGCAACCACTTGGCTTTATGCAATTTTCCACCA 720
 Db 661 AGATCTATATGTCGGCTATTCACAGGGGCAACCACTTGGCTTTATGCAATTTTCCACCA 720
 QY 721 TGGCAGAGCTGGCTCAGAAAATCAAAATGTAATTTTGTCTTACAGCCATAGCCACTGTTA 780
 Db 721 TGGCAGAGCTGGCTCAGAAAATCAAAATGTAATTTTGTCTTACAGCCATAGCCACTGTTA 780
 QY 781 AGCATGAAAAGCCCGGACCAAAATTTTGTCTTACAGCCATAGCCACTGTTA 840
 Db 781 AGCATGAAAAGCCCGGACCAAAATTTTGTCTTACAGCCATAGCCACTGTTA 840
 QY 841 TGTGTGCAAAAAGAAATTTCTGATCAGACCAAGATTTCTCAGACAACTTTGTTATTACC 900
 Db 841 TGTGTGCAAAAAGAAATTTCTGATCAGACCAAGATTTCTCAGACAACTTTGTTATTACC 900
 QY 901 TTTGTGCCAGGATGATTTGATCAGATTTGTAATATATCATGTTTCTGGGTGGAT 960
 Db 901 TTTGTGCCAGGATGATTTGATCAGATTTGTAATATATCATGTTTCTGGGTGGAT 960
 QY 961 TCAACACCAATATGAAATGAGCGGACCAAGTATGATGTCGCCACACTTTGCTG 1020
 Db 961 TCAACACCAATATGAAATGAGCGGACCAAGTATGATGTCGCCACACTTTGCTG 1020
 QY 1021 GAACATCTGTGCAAAATATTTCTACCTGGAGCCAGGAGTGAATTTCTGTGAACTCCGG 1080
 Db 1021 GAACATCTGTGCAAAATATTTCTACCTGGAGCCAGGAGTGAATTTCTGTGAACTCCGG 1080
 QY 1081 CATTTGACTGGGGAGTGAGACCAAAATCTGGAATAATCAATCAAGCACTCTGTAA 1140
 Db 1081 CATTTGACTGGGGAGTGAGACCAAAATCTGGAATAATCAATCAAGCACTCTGTAA 1140
 QY 1141 GGTACAGAGTCAGAGATATGAGGCTCCCTACAGCAATGTGCGACAGAGGTCAGGCTGC 1200
 Db 1141 GGTACAGAGTCAGAGATATGAGGCTCCCTACAGCAATGTGCGACAGAGGTCAGGCTGC 1200
 QY 1201 TTTCAAATCCAGAGAGCTGAAAATGCTGCTCTCTGAGGTGACCAACCTCATCTACATA 1260
 Db 1201 TTTCAAATCCAGAGAGCTGAAAATGCTGCTCTCTGAGGTGACCAACCTCATCTACATA 1260
 QY 1261 AGAATATTTCTGAATGGCTCAGTGGAATTTCACTTGGGGTTTGGATGCTCCTCAGGTA 1320
 Db 1261 AGAATATTTCTGAATGGCTCAGTGGAATTTCACTTGGGGTTTGGATGCTCCTCAGGTA 1320
 QY 1321 TGTACAATGAAATCATCCATCTGATGACAGGAGGAGACCAACCTTTCCAGGGACGCT 1380
 Db 1321 TGTACAATGAAATCATCCATCTGATGACAGGAGGAGACCAACCTTTCCAGGGACGCT 1380
 QY 1381 GTGAGGCGGTATTTGAGGATCTGACATCTGACATCTTAGGACCAACCTCTGAGGGATG 1440
 Db 1381 GTGAGGCGGTATTTGAGGATCTGACATCTGACATCTTAGGACCAACCTCTGAGGGATG 1440
 QY 1441 GGGCTAGGACCCATGAGGAGCAATTTACGAGAGCAGAGACCTTAGTATACATTTTTCAGA 1500
 Db 1441 GGGCTAGGACCCATGAGGAGCAATTTACGAGAGCAGAGACCTTAGTATACATTTTTCAGA 1500
 QY 1501 TTCCCTGCACTTGGCACTAAATCCGACATTTAATTACATTTTCTGTAAATATAA 1560
 Db 1501 TTCCCTGCACTTGGCACTAAATCCGACATTTAATTACATTTTCTGTAAATATAA 1560
 QY 1561 GTACTTATAGGTAATAGAGGTTTGTATGCTTATATATTTCTACCATCTTGAAGGT 1620

Db 1561 GTACTTATAGGTAATAGAGGTTTGTATGCTTATATATTTCTACCATCTTGAAGGT 1620
 QY 1621 AGGTTTACCTGATAGAGCAAAATATCTAGACATTTCTCTATATCATTTAGGTAATCTC 1680
 Db 1621 AGGTTTACCTGATAGAGCAAAATATCTAGACATTTCTCTATATCATTTAGGTAATCTC 1680
 QY 1681 TTTAAACACCTATTTGTTTTTCTATAAGCCATATTTTGGAGCACTAAAGTAAATGGC 1740
 Db 1681 TTTAAACACCTATTTGTTTTTCTATAAGCCATATTTTGGAGCACTAAAGTAAATGGC 1740
 QY 1741 AAATTTGGACAGATATTTGAGTCTGGAGTCTGTGATTTTGTGATTTTGTGACAAAATAA 1800
 Db 1741 AAATTTGGACAGATATTTGAGTCTGGAGTCTGTGATTTTGTGATTTTGTGACAAAATAA 1800
 QY 1801 GCTAGACATTTTACCTTTGTTGCCACAGACATACACACTACCTCAGGAAAGCTGAGCTGC 1860
 Db 1801 GCTAGACATTTTACCTTTGTTGCCACAGACATACACACTACCTCAGGAAAGCTGAGCTGC 1860
 QY 1861 TTTAAGGACAAACAACAATAATCAGTTTACAGTATGAGTAAATCTATGTTAAGCATT 1920
 Db 1861 TTTAAGGACAAACAACAATAATCAGTTTACAGTATGAGTAAATCTATGTTAAGCATT 1920
 QY 1921 CTCAGATAAGGCAAGTTTATATGTTGATCTCAGGGAAGAAAATTTTATAGGATGTT 1980
 Db 1921 CTCAGATAAGGCAAGTTTATATGTTGATCTCAGGGAAGAAAATTTTATAGGATGTT 1980
 QY 1981 ATGAGTTCTCAATAAATGCAATTTCTGCAATTTACATTAATAAAAAAAGGCGG 2040
 Db 1981 ATGAGTTCTCAATAAATGCAATTTCTGCAATTTACATTAATAAAAAAAGGCGG 2040
 QY 2041 CCGC 2044
 Db 2041 CCGC 2044

RESULT 3
 AAD19226
 ID AAD19226 standard; DNA; 2035 BP.
 XX
 AC AAD19226;
 XX
 DT 18-DEC-2001 (first entry)
 XX
 DE Human CG162 (or C59) lipase DNA #2.
 XX
 KW Human; apolipoprotein; lipase; lipoprotein receptor; ALLr; angina; cardiovascular disease; lipid metabolism; myocardial infarction; cerebral ischaemia; arterial thrombosis; thrombolytic; antilipemic; coronary artery thrombosis; cerebral artery thrombosis; stroke; intracardiac thrombosis; gene therapy; cardiovascular; vasodilator; neuroprotectant; cerebroprotective; ds.
 XX
 OS Homo sapiens.
 XX
 FH Key Location/Qualifiers
 FT CDS 189..1418
 FT /*tag= a
 FT /product= "Human CG162 (or C59) lipase protein"
 FT sig_peptide 189..245
 FT /*tag= b
 FT mat_peptide 246..1415
 FT /*tag= c
 FT /product= "Human mature CG162 (or C59) lipase protein"
 XX
 PN WO200179446-A2.
 XX
 PD 25-OCT-2001.
 XX
 PF 16-APR-2001; 2001WO-US12529.
 XX
 PR 14-APR-2000; 2000US-197137P.
 PR 20-JUN-2000; 2000US-0598042.

PR 03-AUG-2000; 2000US-0631451.
 PR 22-SEP-2000; 2000US-0667298.
 PR 17-NOV-2000; 2000US-0714936.
 PA (HYSE-) HYSEQ INC.
 XX Ballinger DG, Loeb D, Montgomery JR, Tang TY, Zhou P, Goodrich R;
 PI Liu C, Asundi V, Zhao QA, Wehrman T, Drmanac RT, Ren F, Qian XB;
 PI Wang D;
 XX WPI: 2001-611724/70.
 DR P-PSDB; A2E11931.
 DR Nucleic acids encoding human apolipoproteins, lipases, and lipoprotein
 XX receptor polypeptides, useful for preventing diagnosing and treating
 PT lipid metabolism disorders, thrombosis and cardiovascular diseases -
 PT Claim 1; Page 183-185; 266pp; English.
 XX The invention relates to polynucleotides encoding proteins CG122, CG179,
 CC CG95, CG121, CG162, CG27, CG153 and CG168 which are related to proteins
 CC involved in lipid metabolism and cardiovascular disease such as human
 CC apolipoproteins, lipases and lipoprotein receptor proteins. These DNA
 CC and protein sequences are useful for treating or preventing disorders
 CC associated with apolipoproteins, lipases and lipoprotein receptor (ALLR)
 CC expression and for treating lipid metabolism, cardiovascular diseases
 CC and thrombosis. Antibodies against these proteins are useful for
 CC determining the presence of or predisposition to a disease associated
 CC with altered levels of these sequences. ALLR polypeptides are also
 CC useful for identifying agents (agonists and antagonists) that bind to
 CC them and cells expressing ALLR proteins are useful for identifying a
 CC therapeutic agent for use in treatment of a pathology related to
 CC aberrant expression or physiological interactions of this polypeptide.
 CC Vectors comprising these DNA and protein sequences are also useful for
 CC producing ALLR proteins. The nucleic acids and polypeptides of the
 CC invention are also useful for the treatment of occlusive cardiovascular
 CC diseases, myocardial infarction, cerebral ischaemia, angina, arterial
 CC thrombosis, coronary artery thrombosis and cerebral artery thrombosis
 CC or intracardiac thrombosis and stroke. The nucleotides of the invention
 CC are used in gene therapy. The present sequence is human CG162 (or C59)
 CC lipase DNA.
 XX
 SQ Sequence 2035 BP; 608 A; 395 C; 456 G; 576 T; 0 other;
 Query Match 97.5%; Score 1993.8; DB 22; Length 2035;
 Best Local Similarity 99.9%; Pred. No. 0;
 Matches 1995; Conservative 0; Mismatches 2; Indels 0; Gaps 0;
 QY 18 GGGAAATTCAGCAGGAGAAATATGCAAGAGATTTTAAACCCCAAAATTCCTTACTTTA 77
 DB 39 GGGAAATTCAGCAGGAGAAATATGCAAGAGATTTTAAACCCCAAAATTCCTTACTTTA 98
 QY 78 GAATTAAGTTTACATTCGAGGAGAAATTAATGACAGATTTGGAACATTTGGAAACC 137
 DB 99 GAATTAAGTTTACATTCGAGGAGAAATTAATGACAGATTTGGAACATTTGGAAACC 158
 QY 138 TTGTCAAGACAGTGGATTTCTCACAAGAAATGGAATGAGCTTCTGATTTCTGGTGGC 197
 DB 159 TTGTCAAGACAGTGGATTTCTCACAAGAAATGGAATGAGCTTCTGATTTCTGGTGGC 218
 QY 198 TATATGTTCCAGAGAAATGTGAATTCAGTACATATGCCAACTAAAGCTGTGGACCCAGAA 257
 DB 219 TATATGTTCCAGAGAAATGTGAATTCAGTACATATGCCAACTAAAGCTGTGGACCCAGAA 278
 QY 258 GCATTCATGATATTAAGTGAATATCAACATCAAGGCTATCCCTGTGAGGAATATGAA 317
 DB 279 GCATTCATGATATTAAGTGAATATCAACATCAAGGCTATCCCTGTGAGGAATATGAA 338
 QY 318 GTCCGAATCTGAAGATGGGTATATCTTCTGTGAACAGATTCCTCGAGGCTTAGTGCAA 377
 DB 339 GTCCGAATCTGAAGATGGGTATATCTTCTGTGAACAGATTCCTCGAGGCTTAGTGCAA 398
 QY 378 CCTAAGAAGACAGGTTCCAGGCGCTGTGGTGTACTGACAGATGGCGCTTAGTGGAGTGCT 437

DB 399 CCTAAGAGACAGGTTCCAGGCGCTGTGGTGTACTGACAGATGGCTAGTTCGAGGTGCT 458
 QY 438 AGCAACTGGATTTCCAACTGCGCCAAATAGCTGGGCTTCAATCTGCGCATGCTGTGT 497
 DB 459 AGCAACTGGATTTCCAACTGCGCCAAATAGCTGGGCTTCAATCTGCGCATGCTGTGT 518
 QY 498 TTTGACGTGTGGATGGGGAACAGCAGGAGGAAACGCTGTCTCGAAACACACAGACATCTC 557
 DB 519 TTTGACGTGTGGATGGGGAACAGCAGGAGGAAACGCTGTCTCGAAACACACAGACATCTC 578
 QY 558 TCCATAGACCAAGATGATGTTCTGGGCTTTTCAGTTATGATGAGATGCTAGTGTGACCTTT 617
 DB 579 TCCATAGACCAAGATGATGTTCTGGGCTTTTCAGTTATGATGAGATGCTAGTGTGACCTTT 638
 QY 618 CTTGACGTGATAACTTTATTTTGCAGAAACGCGCCAGGAAAGATCTATTTATCTGCGC 677
 DB 639 CTTGACGTGATAACTTTATTTTGCAGAAACGCGCCAGGAAAGATCTATTTATCTGCGC 698
 QY 678 TATTCACAGGCGCAACCACTAGGCTTTTATGATTTTCCACCATGCGCAGCTGGCTCAG 737
 DB 699 TATTCACAGGCGCAACCACTAGGCTTTTATGATTTTCCACCATGCGCAGCTGGCTCAG 758
 QY 738 AAAATCAAAATGATTTTCTTTAGCACCATAGCAGCTGTTAAGCATGCAAAAAGCCCC 797
 DB 759 AAAATCAAAATGATTTTCTTTAGCACCATAGCAGCTGTTAAGCATGCAAAAAGCCCC 818
 QY 798 GGGACCAAAATTTTGTGTGTCGAGATATGATGATCAAGGATTTGTTGGCAAAAAGAA 857
 DB 819 GGGACCAAAATTTTGTGTGTCGAGATATGATGATCAAGGATTTGTTGGCAAAAAGAA 878
 QY 858 TTTCTGTATCAGACCAAGATTTCTCAGACAACTGTTATTTTACCTTTGTGGCCAGGTGAT 917
 DB 879 TTTCTGTATCAGACCAAGATTTCTCAGACAACTGTTATTTTACCTTTGTGGCCAGGTGAT 938
 QY 918 CTTGTACAGATTTGTAGTAATATCATGTTACTCTTGGGTGATTCACACCAACAAATATG 977
 DB 939 CTTGTACAGATTTGTAGTAATATCATGTTACTCTTGGGTGATTCACACCAACAAATATG 998
 QY 978 AACATGAGCGGAGCAAGTGTATATGCTGCCACACTCTTGTCTGGAACTCTGTGCAAAAT 1037
 DB 999 AACATGAGCGGAGCAAGTGTATATGCTGCCACACTCTTGTCTGGAACTCTGTGCAAAAT 1058
 QY 1038 ATTCTACATGAGCGGAGCAAGTGTATGCTGTGAACTCTGGGCAATTTGACTGGGGAGT 1097
 DB 1059 ATTCTACATGAGCGGAGCAAGTGTATGCTGTGAACTCTGGGCAATTTGACTGGGGAGT 1118
 QY 1098 GAGACCAAAATCTGGAAAAATGCAATCAGCAGCAACTCTCTGTAAGGTACAGAGTCAAGAT 1157
 DB 1119 GAGACCAAAATCTGGAAAAATGCAATCAGCAGCAACTCTCTGTAAGGTACAGAGTCAAGAT 1178
 QY 1158 ATGACGGTCCCTACAGCAATGTGGACAGAGGTCAAGACTGGCTTTCAAATCCAGAAAGAC 1217
 DB 1179 ATGACGGTCCCTACAGCAATGTGGACAGAGGTCAAGACTGGCTTTCAAATCCAGAAAGAC 1238
 QY 1218 GTGAAAAATCTGCTCTCTGAGGTGAGCCACTCTCTACCAATGAATATTTCTGAAATGG 1277
 DB 1239 GTGAAAAATCTGCTCTCTGAGGTGAGCCACTCTCTACCAATGAATATTTCTGAAATGG 1298
 QY 1278 GCTCAGCTGGATTTCACTCTGGGCTTTGGATGCTCTCCCTGATGTATGATGAAATCATC 1337
 DB 1299 GCTCAGTGGATTTCACTCTGGGCTTTGGATGCTCTCCCTGATGTATGATGAAATCATC 1358
 QY 1338 CATCTGATCAGCAGGAGGAGCAACCTTTCCAGGAGCGGTGTGAGGCCGTATTTGGA 1397
 DB 1359 CATCTGATCAGCAGGAGGAGCAACCTTTCCAGGAGCGGTGTGAGGCCGTATTTGGA 1418
 QY 1398 AGCATCTGACACTGACATCTTAGGACAACTCTCTGAGGAGTGGGCTTAGGCCCATGAA 1457
 DB 1419 AGCATCTGACACTGACATCTTAGGACAACTCTCTGAGGAGTGGGCTTAGGCCCATGAA 1478
 QY 1458 GGCAGAAATTCGAGAGACAGACCTAGTATATATTTTTCAGATTCCTCTGCACTTGGCAC 1517

Db 1479 GGCAGAAATATGAGAGACAGACACCTAGTATACATTTTTCAGATCCCTGCACCTGGCAC 1538
 QY 1518 TAAATCGACACTTACATTTACATTTTTCCTGTAATTAAGTACTTATTAGTAAT 1577
 Db 1539 TAAATCGACACTTACATTTACATTTTTCCTGTAATTAAGTACTTATTAGTAAT 1598
 QY 1578 AGAGGTTTTGTATGCTATTATATATTTCTACCATCTTGAAGGCTAGTTTACCTGATAGC 1637
 Db 1599 AGAGGTTTTGTATGCTATTATATATTTCTACCATCTTGAAGGCTAGTTTACCTGATAGC 1658
 QY 1638 CAGAAATATCTAGACATTTCTATATATCTAGGTAATCTCTTTAAACACCTATTGT 1697
 Db 1659 CAGAAATATCTAGACATTTCTATATATCTAGGTAATCTCTTTAAACACCTATTGT 1718
 QY 1698 TTTTCTATAGCCATATTTTTCGACACTTAAAGTAAATGCAATTCGGGACAGATAT 1757
 Db 1719 TTTTCTATAGCCATATTTTTCGACACTTAAAGTAAATGCAATTCGGGACAGATAT 1778
 QY 1758 GAGGCTGGAGTCTGGATTTATTTGTGACTTTTGAACAAATAGCTAGACATTTTCACT 1817
 Db 1779 GAGGCTGGAGTCTGGATTTATTTGTGACTTTTGAACAAATAGCTAGACATTTTCACT 1838
 QY 1818 TGTTCACAGACACATACACTACCTCCAGAGCTGAGCTGCTTTAAGGACACACAA 1877
 Db 1839 TGTTCACAGACACATACACTACCTCCAGAGCTGAGCTGCTTTAAGGACACACAA 1898
 QY 1878 CAAATCAGTGTACAGTATGATGAATCTATGTTAAGCAATCTCAGAAATAGGCCAAG 1937
 Db 1899 CAAATCAGTGTACAGTATGATGAATCTATGTTAAGCAATCTCAGAAATAGGCCAAG 1958
 QY 1938 TTTTATAGTGCATCTCAGGAGAAATTTTATAGATGTTTATAGATTTTCAATATA 1997
 Db 1959 TTTTATAGTGCATCTCAGGAGAAATTTTATAGATGTTTATAGATTTTCAATATA 2018
 QY 1998 TGCATTTCTGCATACAT 2014
 Db 2019 TGCATTTCTGCATACAT 2035

RESULT 4

AA019220
 ID AAD19220 standard; DNA; 1384 BP.
 XX
 AC AAD19220;
 XX
 DT 18-DEC-2001 (first entry)
 XX
 DE Human CG162 (or C59) lipase DNA #1.
 XX
 KW Human; apolipoprotein; lipase; lipoprotein receptor; ALR; angina;
 KW cardiovascular disease; lipid metabolism; myocardial infarction;
 KW cerebral ischaemia; arterial thrombosis; thrombolytic; antilipemic;
 KW coronary artery thrombosis; cerebral artery thrombosis; stroke;
 KW intracardiac thrombosis; gene therapy; cardiovascular; vasodilator;
 KW neuroprotectant; cerebroprotective; ds.
 XX
 OS Homo sapiens.
 XX
 FN WO200179446-A2.
 XX
 PD 25-OCT-2001.
 XX
 PF 16-APR-2001; 2001WO-US12529.
 XX
 PR 14-APR-2000; 2000US-197137P.
 PR 20-JUN-2000; 2000US-0598042.
 PR 03-AUG-2000; 2000US-0631451.
 PR 22-SEP-2000; 2000US-0667298.
 PR 17-NOV-2000; 2000US-0714936.
 XX
 PA (HYSE-) HYSEQ INC.
 XX
 PI Ballinger DG, Loeb D, Montgomery JR, Tang TY, Zhou P, Goodrich R;

PI Liu C, Asundi V, Zhao QA, Wehrman T, Drmanac RT, Ren F, Qian XB;
 PI Wang D;
 XX
 DR WPI; 2001-611724/70.
 XX
 PT Nucleic acids encoding human apolipoproteins, lipases, and lipoprotein
 PT receptor polypeptides, useful for preventing diagnosing and treating
 PT lipid metabolism disorders, thrombosis and cardiovascular diseases -
 XX
 PS Claim 1; Page 160-161; 266pp; English.
 XX
 CC The invention relates to polynucleotides encoding proteins CG122, CG179,
 CC CG95, CG121, CG162, CG27, CG153 and CG168 which are related to proteins
 CC involved in lipid metabolism and cardiovascular disease such as human
 CC apolipoproteins, lipases and lipoprotein receptor proteins. These DNA
 CC and protein sequences are useful for treating or preventing disorders
 CC associated with apolipoproteins, lipases and lipoprotein receptor (ALR)
 CC expression and for treating lipid metabolism, cardiovascular diseases
 CC and thrombosis. Antibodies against these proteins are useful for
 CC determining the presence of or predisposition to a disease associated
 CC with altered levels of these sequences. ALR polypeptides are also
 CC useful for identifying agents (agonists and antagonists) that bind to
 CC them and cells expressing ALR proteins are useful for identifying a
 CC therapeutic agent for use in treatment of a pathology related to
 CC aberrant expression or physiological interactions of this polypeptide.
 CC Vectors comprising these DNA and protein sequences are also useful for
 CC producing ALR proteins. The nucleic acids and polypeptides of the
 CC invention are also useful for the treatment of occlusive cardiovascular
 CC diseases, myocardial infarction, cerebral ischaemia, angina, arterial
 CC thrombosis, coronary artery thrombosis and cerebral artery thrombosis
 CC or intracardiac thrombosis and stroke. The nucleotides of the invention
 CC are used in gene therapy. The present sequence is human CG162 (or C59)
 CC Lipase DNA.
 XX
 SQ Sequence 1384 BP; 404 A; 281 C; 328 G; 371 T; 0 other;
 Query Match 65.7%; Score 1343.4; DB 22; Length 1384;
 Best Local Similarity 99.9%; Pred. No. 0;
 Matches 1344; Conservative 0; Mismatches 1; Indels 0; Gaps 0;
 QY 18 GGGAAATTCAGCAGGAAATATATGGAAGAGTTTAAACCCACAAATCTTCTTACTTTA 77
 Db 39 GGGAAATTCAGCAGGAAATATATGGAAGAGTTTAAACCCACAAATCTTCTTACTTTA 98
 QY 78 GAATAGTTGTACATTTGGCAGGAAATTAATGAGATGTTGGACATGTTGGAAC 137
 Db 99 GAATAGTTGTACATTTGGCAGGAAATTAATGAGATGTTGGACATGTTGGAAC 158
 QY 138 TTGTCAACACAGTGGATTTCTCACACAGAAATGGAATGTTGTTCTGTTGGCG 197
 Db 159 TTGTCAACACAGTGGATTTCTCACACAGAAATGGAATGTTGTTCTGTTGGCG 218
 QY 198 TATATGTTCCAGAGAAATGGAATTCAGTACATATGCAATTAAGCTGTGGACCCAGAA 257
 Db 219 TATATGTTCCAGAGAAATGGAATTCAGTACATATGCAATTAAGCTGTGGACCCAGAA 278
 QY 258 GCATTTCATGAATATTAGTGAATCATCAACATCAAGCTATCCCTGTGAGGAATATGAA 317
 Db 279 GCATTTCATGAATATTAGTGAATCATCAACATCAAGCTATCCCTGTGAGGAATATGAA 338
 QY 318 GTCCCAACTGAGATGGGTATATCTTTTCTTTAAACAGGATTCCTCGAGGCTAGTGCA 377
 Db 339 GTCCCAACTGAGATGGGTATATCTTTTCTTTAAACAGGATTCCTCGAGGCTAGTGCA 398
 QY 378 CCTAAGAGACAGGTTCCAGGCTGTGTTTACTTGCAGCATGCTAGTTGGAGGTGCT 437
 Db 399 CCTAAGAGACAGGTTCCAGGCTGTGTTTACTTGCAGCATGCTAGTTGGAGGTGCT 458
 QY 438 AGCAACTGGATTTCCAACTCCGCAACATAGCTGCTTCAATCTGCGCAGATGCTGGT 497
 Db 459 AGCAACTGGATTTCCAACTCCGCAACATAGCTGCTTCAATCTGCGCAGATGCTGGT 518
 QY 498 TTTCAGTGTGATGGGAAACAGCAGGAGAAACGCTGCTCGGAAACACACACACTC 557

Db 519 TTTGACGTGGTGGGAAACAGCAGGGGAAACGCGTGTCTCGAAACACAGACACTC 578
 Qy 558 TCCATAGACCAAGATGAGTCTGGGCTTCAGTATATGATGAGATGCTAGGTTGACCTT 617
 Db 579 TCCATAGACCAAGATGAGTCTGGGCTTCAGTATATGATGAGATGCTAGGTTGACCTT 638
 Qy 618 CCTGACGTATAACTTTATTTTGCAGAAACCGGGCCAGGAAAGATCTATATGTCGGC 677
 Db 639 CCTGACGTATAACTTTATTTTGCAGAAACCGGGCCAGGAAAGATCTATATGTCGGC 698
 Qy 678 TATTACAGGGCACCACATGGGCTTTATTTGCAATTTTCCACCATGCCAGAGCTGGCTCAG 737
 Db 699 TATTACAGGGCACCACATGGGCTTTATTTGCAATTTTCCACCATGCCAGAGCTGGCTCAG 758
 Qy 738 AAAATCAAATGATTTTGTCTTAGCACCCATGACCACTGTAAAGCATGCAAAAAGCCGC 797
 Db 759 AAAATCAAATGATTTTGTCTTAGCACCCATGACCACTGTAAAGCATGCAAAAAGCCGC 818
 Qy 798 GGGACCAAATTTTGTCTGCCAGATATGATGATCAAGGATGTTTGGCAAAAAGAA 857
 Db 819 GGGACCAAATTTTGTCTGCCAGATATGATGATCAAGGATGTTTGGCAAAAAGAA 878
 Qy 858 TTTCTGATCAGACCACTTCTCAGACAACTTGTATTTTACCTTTGTCGCCAGGTGATT 917
 Db 879 TTTCTGATCAGACCACTTCTCAGACAACTTGTATTTTACCTTTGTCGCCAGGTGATT 938
 Qy 918 CTTGATCAGATTTTGTAGTATATCATGTTTCTGGGTGGATTCAACACCAATATG 977
 Db 939 CTTGATCAGATTTTGTAGTATATCATGTTTCTGGGTGGATTCAACACCAATATG 998
 Qy 978 AACATGACCGCAGCAAGTATATGCTGCCACACTTCTGCGAACATCTGTGCAAAAT 1037
 Db 999 AACATGACCGCAGCAAGTATATGCTGCCACACTTCTGCGAACATCTGTGCAAAAT 1058
 Qy 1038 ATTCTACACTGGACCCAGGAGTGAATCTCTGGTGAATCTCGGGCAATTTGATCGGGGAGT 1097
 Db 1059 ATTCTACACTGGACCCAGGAGTGAATCTCTGGTGAATCTCGGGCAATTTGATCGGGGAGT 1118
 Qy 1098 GAGACCAAATCTGGAAATGCAATCAGCCAACTCTCTGTAAGGTACAGATCAGAGAT 1157
 Db 1119 GAGACCAAATCTGGAAATGCAATCAGCCAACTCTCTGTAAGGTACAGATCAGAGAT 1178
 Qy 1158 ATGACGGTCCCTCAGCAATCTGACAGAGGAGTTCAGACTGGCTTCCAAATCCAGAGAC 1217
 Db 1179 ATGACGGTCCCTCAGCAATCTGACAGAGGAGTTCAGACTGGCTTCCAAATCCAGAGAC 1238
 Qy 1218 GTGAAATGCTGCTCTCTGAGGTGACCAACCTCATCTACCAATGAATATTTCTGAAATGG 1277
 Db 1239 GTGAAATGCTGCTCTCTGAGGTGACCAACCTCATCTACCAATGAATATTTCTGAAATGG 1298
 Qy 1278 GCTCAGGTGATTTTCACTCTGGGTTTGGATGCTCTCAGCGTATGTACATGAATCATC 1337
 Db 1299 GCTCATGTTGGATTTTCACTCTGGGTTTGGATGCTCTCAGCGTATGTACATGAATCATC 1358
 Qy 1338 CATCTGATGACAGGAGGAGACCA 1362
 Db 1359 CATCTGATGACAGGAGGAGACCA 1383

RESULT 5
 AAF45132
 ID AAF45132 standard; cDNA; 1269 BP.
 XX
 XX AAF45132;
 AC
 DT 30-MAR-2001 (first entry)
 XX
 XX Human TANGO 294 ORF.
 XX
 KW Gene therapy; TANGO protein; INTERCEPT protein; neurological disorder;
 KW central nervous system; focal brain disorder; bipolar affective disorder;
 KW global-diffuse cerebral disorder; cerebrovascular; Alzheimer's disease;

senile dementia; Huntington's disease; amyotrophic lateral sclerosis;
 Parkinson's; Gilles de la Tourette's syndrome; hypertension; sleep;
 neuropsychiatric; psychoactive substance use; anxiety; ss.
 OS Homo sapiens.
 FN WO200077239-A2.
 XX 21-DEC-2000.
 XX 24-MAY-2000; 2000WO-US14858.
 XX 14-JUN-1999; 99US-0333159.
 XX (MILL-) MILLENNIUM PHARM INC.
 XX McCarthy SA, Fraser CC, Sharp JD, Barnes TM;
 DR WPI: 2001-032313/04.
 DR P-PSDB; AAB66065.
 XX TANGO and INTERCEPT nucleic acids, proteins, and antibodies, useful for
 screening assays and diagnostic assays and for the treatment of
 neurological diseases such as Alzheimer's, Parkinson's and Huntington's
 disease.
 Claim 1; Fig 6; 359pp; English.
 The present invention relates to TANGO or INTERCEPT proteins and coding
 sequences (see AAF45121-F45136 and AAF45138-F45139 and AAB66031-B66057,
 AAB66064-B66083 and AAB66085). The TANGO/INTERCEPT proteins and coding
 sequences are useful for the treatment of neurological disorders such as
 central nervous system (CNS) disorders, CNS-related disorders, focal
 brain disorders, global-diffuse cerebral disorders and other
 neurological and cerebrovascular disorders. The CNS disorders include
 Alzheimer's disease, senile dementia, Huntington's disease, amyotrophic
 lateral sclerosis, Parkinson's, Gilles de la Tourette's syndrome,
 autonomic function disorders such as hypertension and sleep disorders,
 neuropsychiatric disorders, psychoactive substance use disorders,
 anxiety, and bipolar affective disorder.

Query Match 62.1%; Score 1269; DB 22; Length 1269;
 Best Local Similarity 100.0%; Pred. No. 0;
 Matches 1269; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 Qy 126 ATGTTGGAAACCTTCTCAAGACAGTGGATTCTCTCACACAGAAATGGAATGTGGCTTCTG 185
 Db 1 ATGTTGGAAACCTTCTCAAGACAGTGGATTCTCTCACACAGAAATGGAATGTGGCTTCTG 60
 Qy 186 ATTCCTGGTGGGTATATGTTCCAGAGAAATGTGAATTCAGTACATATGCCAACTAAGCT 245
 Db 61 ATTCCTGGTGGGTATATGTTCCAGAGAAATGTGAATTCAGTACATATGCCAACTAAGCT 120
 Qy 246 GTGGACCCAGAGCAATTCATGAATATTAGTGAATATCATCCAACTAAGGCTATCCCTGT 305
 Db 121 GTGGACCCAGAGCAATTCATGAATATTAGTGAATATCATCCAACTAAGGCTATCCCTGT 180
 Qy 306 GAGGAATATGAAGTGGCACTGAGATGGGTATATCTCTTCTGTTAAACAGGATTCCTCGA 365
 Db 181 GAGGAATATGAAGTGGCACTGAGATGGGTATATCTCTTCTGTTAAACAGGATTCCTCGA 240
 Qy 366 GGCCTAGTGCACACCTTAAGAGACAGGTTCCAGGCTGTGGTGTACTTACTCAGCATGGCCTA 425
 Db 241 GGCCTAGTGCACACCTTAAGAGACAGGTTCCAGGCTGTGGTGTACTTACTCAGCATGGCCTA 300
 Qy 426 GTTGGAGGTGCTAGCAACTGGATTTCCAACTGCCCAACAATAGCTGGGCTTCAATCTG 485
 Db 301 GTTGGAGGTGCTAGCAACTGGATTTCCAACTGCCCAACAATAGCTGGGCTTCAATCTG 360
 Qy 486 GCAGATGCTGGTGTGAGCTGTGGATGGGACAGCGGGAACGCTGTCTCGAAAA 545

Db 361 GCAGATGCTGTTTGGATGCTGGGAAACAGACAGGGGAAACGCTGCTCGAAA 420
 Qy 546 CACAAGACACTCTCCATAGACCAAGATGATGCTGGCTTTCAGTTATGATGAGATGCT 605
 Db 421 CACAAGACACTCTCCATAGACCAAGATGATGCTGGCTTTCAGTTATGATGAGATGCT 480
 Qy 606 AGCTTTGACCTTCTGACGATGATAAATTTTTCGAGAAACGGGCGGCAAGGATC 665
 Db 481 AGCTTTGACCTTCTGACGATGATAAATTTTTCGAGAAACGGGCGGCAAGGATC 540
 Qy 666 TATTATGTCGGCTATTTCACAGGACCAACCATGGGCTTTATGCAATTTTCACACGCA 725
 Db 541 TATTATGTCGGCTATTTCACAGGACCAACCATGGGCTTTATGCAATTTTCACACGCA 600
 Qy 726 GAGCTGCTCAGAAATCAAAATGATTTTTCGACACCATAGCCACTGCTTAAGCAT 785
 Db 601 GAGCTGCTCAGAAATCAAAATGATTTTTCGACACCATAGCCACTGCTTAAGCAT 660
 Qy 786 GCAAAAGCCCGGACCAAAATTTTTCGTCGCCAGATATGATGATCAAGGATGTTT 845
 Db 661 GCAAAAGCCCGGACCAAAATTTTTCGTCGCCAGATATGATGATCAAGGATGTTT 720
 Qy 846 GGCAGGATGATTTCTGATCAGATTTTGTAGTAATATCATGTTACTTCTGGGTGATCAAC 965
 Db 781 GGCAGGATGATTTCTGATCAGATTTTGTAGTAATATCATGTTACTTCTGGGTGATCAAC 840
 Qy 966 ACCAACAATATGAACATGAGCCGAGCAAGTGTATGCTGCCACACACTTCTGCTGAA 1025
 Db 841 ACCAACAATATGAACATGAGCCGAGCAAGTGTATGCTGCCACACACTTCTGCTGAA 900
 Qy 1026 TCTGTGCAAAATATTTCTACACTGAGCCGAGCAAGTGTATGCTGCCACACACTTCTGCTGAA 1085
 Db 901 TCTGTGCAAAATATTTCTACACTGAGCCGAGCAAGTGTATGCTGCCACACACTTCTGCTGAA 960
 Qy 1086 GACTGGGGAGTCAGACCAAAATCTGGAAAATGCAATCAGCCAACTCTCTGTAAGGTAC 1145
 Db 961 GACTGGGGAGTCAGACCAAAATCTGGAAAATGCAATCAGCCAACTCTCTGTAAGGTAC 1020
 Qy 1146 AGAGTCAGATATGAGGTCCTTACAGCAATGTCAGAGGAGGTCAGAGCTGCTTCA 1205
 Db 1021 AGAGTCAGATATGAGGTCCTTACAGCAATGTCAGAGGAGGTCAGAGCTGCTTCA 1080
 Qy 1206 AATCCAGACAGCTGAAAATGCTGCTCTCTGAGGTGACCAACCTCATCTACCATAGAT 1265
 Db 1081 AATCCAGACAGCTGAAAATGCTGCTCTCTGAGGTGACCAACCTCATCTACCATAGAT 1140
 Qy 1266 ATTCTGTAATGGGCTCAGTGGATTTTCATCTGGGTTTGGATGCTCTCAGGTATGAT 1325
 Db 1141 ATTCTGTAATGGGCTCAGTGGATTTTCATCTGGGTTTGGATGCTCTCAGGTATGAT 1200
 Qy 1326 AATGAATCATCATCTGATGTCAGAGGAGGAGACCAACCTTTCCAGGAGCGGTGTAG 1385
 Db 1201 AATGAATCATCATCTGATGTCAGAGGAGGAGACCAACCTTTCCAGGAGCGGTGTAG 1260
 Qy 1386 GCGGTATTG 1394
 Db 1261 GCGGTATTG 1269

RESULT 6
 ID ABK86569 standard; cDNA; 1210 BP.
 XX
 AC ABK86569;
 XX
 DT 24-SEP-2002 (first entry)
 XX
 DE Human cDNA encoding lysosomal acid lipase #1.
 XX

Human; ss; gene; lysosomal acid lipase; lipid malabsorption illness;
 cystic fibrosis; alcoholism; heart disease; heart attack;
 Wolman disease; cholesterol ester storage disease; brain injury;
 mood disorder; anxiety disorder; thought disorder; volition disorder;
 sleep disorder; neurogenic disorder; myopathic disorder; COPD;
 obesity; cancer; chronic obstructive pulmonary disease; diabetes;
 cardiovascular disorder; Alzheimer's disease; Parkinson's disease;
 anorexia; osteoarthritis; central nervous system disorder;
 peripheral nervous system disorder.
 Homo sapiens.
 Key Location/Qualifiers
 CDS 1..1197
 /*tag= a
 /product= "Lysosomal acid lipase #1"
 WO200236731-A2.
 10-MAY-2002.
 30-OCT-2001; 2001WO-EPI2518.
 31-OCT-2000; 2000US-244170P.
 29-MAY-2001; 2001US-293516P.
 (FARB) BAYER AG.
 Xiao Y;
 WPI; 2002-519248/55.
 P-PSDB; AAU99164.
 Novel human lysosomal acid lipase polypeptide, useful for treating
 cancer, diabetes, obesity, chronic obstructive pulmonary disease,
 peripheral or central nervous system disorder or cardiovascular
 disorder
 Claim 1; Fig 1; 126pp; English.
 The invention relates to a purified human lysosomal acid lipase
 polypeptide. Also included are the polynucleotide encoding the
 lipase (or its fragment, derivative, allele or sequence at least 60%
 identical to it), vectors, host cells, a reagent (e.g. an antisense
 oligonucleotide), which binds to the lipase or polynucleotide (used for
 detection and modulating/reducing the lipase activity) and an anti-lipase
 antibody. The lipase and polynucleotide are useful for identifying
 therapeutic agents that either increase or decrease the lipase activity.
 The identified agent, the lipase and polynucleotide are useful for
 treatment of a disease such as lipid malabsorption illness,
 cystic fibrosis, alcoholism, heart disease, heart attack,
 Wolman disease, cholesterol ester storage disease, brain injury,
 mood disorder, anxiety disorder, thought disorder, volition disorder,
 sleep disorder, neurogenic disorder, myopathic disorder, obesity, cancer,
 chronic obstructive pulmonary disease (COPD), diabetes,
 cardiovascular disorder, Alzheimer's disease, Parkinson's disease,
 anorexia, osteoarthritis, a central nervous system disorder and
 a peripheral nervous system disorder. The present sequence is the
 cDNA encoding human lysosomal lipase #1.
 Query Match 55.7%; Score 1138; DB 24; Length 1210;
 Best Local Similarity 100.0%; Pred. No. 1.5e-281;
 Matches 1138; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 Qy 273 AGTGAATCATCCAAATCATCAAGGCTATCCCTGTGAGGAATATGAGTCGAATGAGAT 332
 Db 73 AGTGAATCATCCAAATCATCAAGGCTATCCCTGTGAGGAATATGAGTCGAATGAGAT 132
 Qy 333 GGGTATATCTCTTCTGTTTACAGGATTCCTCGAGGCTAGTCAACCTTGAAGACAGGT 392
 Db 133 GGGTATATCTCTTCTGTTTACAGGATTCCTCGAGGCTAGTCAACCTTGAAGACAGGT 192

QY 393 TCCAGGCTGTGGTGTCTTACTGTCAGCATGGCTTAGTTGGAGGTGCTAGCAACTGGATTTC 452
 Db 193 TCCAGGCTGTGGTGTCTTACTGTCAGCATGGCTTAGTTGGAGGTGCTAGCAACTGGATTTC 252
 QY 453 AACCTGCCCCAAGTAAAGCTGGCTTCAATCTGGCAGATGCTGGTTTGGAGTGTGGATG 512
 Db 253 AACCTGCCCCAAGTAAAGCTGGCTTCAATCTGGCAGATGCTGGTTTGGAGTGTGGATG 312
 QY 513 GGGAAACAGCAGGGGAAAGCGCTGGTCTCGAAACACAAAGACACTCTCCATAGACCAAGAT 572
 Db 313 GGGAAACAGCAGGGGAAAGCGCTGGTCTCGAAACACAAAGACACTCTCCATAGACCAAGAT 372
 QY 573 GAGTTCTGGGCTTTCAGTTATGATGAGATGGCTTAGCTTTGACCTTCTGCGAGTAAAC 632
 Db 373 GAGTTCTGGGCTTTCAGTTATGATGAGATGGCTTAGCTTTGACCTTCTGCGAGTAAAC 432
 QY 633 TTTATTTTGCAGAAAACGGCCAGAAAGATCTATTATGTGGCTATTCTACAGGGCACC 692
 Db 433 TTTATTTTGCAGAAAACGGCCAGAAAGATCTATTATGTGGCTATTCTACAGGGCACC 492
 QY 693 ACCATGGGCTTTATTTGCAATTTTCCACCATGCGCAGAGTGGCTCAGAAAATCAAAATGAT 752
 Db 493 ACCATGGGCTTTATTTGCAATTTTCCACCATGCGCAGAGTGGCTCAGAAAATCAAAATGAT 552
 QY 753 TTTGCTTTAGCACCCATAGCCACTGTTAAGCATGCAAAAGCCCGGACCAAAATTTTG 812
 Db 553 TTTGCTTTAGCACCCATAGCCACTGTTAAGCATGCAAAAGCCCGGACCAAAATTTTG 612
 QY 813 TTGCTGCCAGATPATGATGATCAAGGGATTTGTTGGCAAAAAGAAATTTCTGTATCAGACC 872
 Db 613 TTGCTGCCAGATPATGATGATCAAGGGATTTGTTGGCAAAAAGAAATTTCTGTATCAGACC 672
 QY 873 AGATTTCTCAGCAACTGTTATTTACCTTTGTCGCGAGGTGATTTCTTGATCAGATTGT 932
 Db 673 AGATTTCTCAGCAACTGTTATTTACCTTTGTCGCGAGGTGATTTCTTGATCAGATTGT 732
 QY 933 AGTAATATCATGTTACTTCTGGGTGGATTCAACACCAAAATATGAACATGAGCCGAGCA 992
 Db 733 AGTAATATCATGTTACTTCTGGGTGGATTCAACACCAAAATATGAACATGAGCCGAGCA 792
 QY 993 AGTGTATATGTCGCCACACTTCTGTCGGAACATCTGTGCAAAATATTCTACACTGAGC 1052
 Db 793 AGTGTATATGTCGCCACACTTCTGTCGGAACATCTGTGCAAAATATTCTACACTGAGC 852
 QY 1053 CAGGAGTGAATTTCTGGTGAATCTCGGGCAATTTGACTGGGGAGTGAGACCAAAATCTG 1112
 Db 853 CAGGAGTGAATTTCTGGTGAATCTCGGGCAATTTGACTGGGGAGTGAGACCAAAATCTG 912
 QY 1113 GAAATATGCAATCAGCCAACTCTCTGAAGGTACAGAGTCAGAGATATGACGGTCCCTACA 1172
 Db 913 GAAATATGCAATCAGCCAACTCTCTGAAGGTACAGAGTCAGAGATATGACGGTCCCTACA 972
 QY 1173 GCAATGTGACAGGAGGTGAGGCTTTCATAATCCAGAGAGCTGAAATCTGCTC 1232
 Db 973 GCAATGTGACAGGAGGTGAGGCTTTCATAATCCAGAGAGCTGAAATCTGCTC 1032
 QY 1233 TCTGAGGTGACCAACTCATCTACCAATGAATATTTCTGAAATGGGCTCAGCTGGATTTC 1292
 Db 1033 TCTGAGGTGACCAACTCATCTACCAATGAATATTTCTGAAATGGGCTCAGCTGGATTTC 1092
 QY 1293 ATCTGGGTTTGGATGCTCTCACCGTATGTACATGAATCAATCCATCTGATCAGCAG 1352
 Db 1093 ATCTGGGTTTGGATGCTCTCACCGTATGTACATGAATCAATCCATCTGATCAGCAG 1152
 QY 1353 GAGGAGACCAACTTTTCCAGGGACGGTGTGAGGCCGTATTGTGAAGCATCTGACACT 1410
 Db 1153 GAGGAGACCAACTTTTCCAGGGACGGTGTGAGGCCGTATTGTGAAGCATCTGACACT 1210

RESULT 7
 ABS55356
 ID ABS55356 standard; cdna; 1206 BP.

XX AC ABS55356;
 XX DT 07-JAN-2003 (first entry)
 XX DE cDNA encoding human lipase protein.
 XX KW Human; lipase; cancer; colon; kidney; skin; brain; testis; chromosome 10;
 XX KW Burkitt's lymphoma; cytostatic; gene therapy; lipase-inhibitor; gene; ss.
 XX OS Homo sapiens.
 XX FH Key Location/Qualifiers
 XX FT 5'UTR 1..3
 XX FT FT /*tag= a
 XX FT CDS 4..1200
 XX FT FT /*tag= b
 XX FT FT /*product= "Human lipase protein"
 XX FT 3'UTR 1201..1206
 XX FT FT /*tag= c
 XX XX WO200274975-A2.
 XX XX 26-SBP-2002.
 XX XX 18-MAR-2002; 2002WO-US08035.
 XX XX 20-MAR-2001; 2001US-0811825.
 XX XX (PEKE) PE CORP NY.
 XX XX Yan C, Di Francesco V, Beasley EM;
 XX XX WPI; 2002-750560/81.
 XX XX P-PSDB; ABG71000.
 XX XX New isolated human lipase peptides and encoding nucleic acids, useful
 XX XX for diagnosing and treating disorders mediated by human lipase
 XX XX proteins, e.g. cancer of the colon, kidneys, skin and testis, and
 XX XX Burkitt's lymphoma -
 XX XX Claim 4; Fig 1; 79pp; English.
 XX XX The present invention relates to a new lipase peptide. The methods and
 XX XX compositions of the present invention are useful for diagnosing and
 XX XX treating disorders mediated by the human lipase protein, such as cancer
 XX XX of the colon, kidneys, skin, brain and testis, and Burkitt's lymphoma.
 XX XX The present nucleic acid sequence represents the human lipase gene
 XX XX located on chromosome 10. This sequence encodes the human lipase protein
 XX XX of the invention.
 XX XX Sequence 1206 BP; 339 A; 260 C; 295 G; 312 T; 0 other;
 XX XX
 XX XX Query Match 55.1%; Score 1126; DB 24; Length 1206;
 XX XX Best Local Similarity 100.0%; Fred. No. 1.8e-278;
 XX XX Matches 1126; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 QY 273 AGTGAATCATCCAAATCAAGCTATCCCTGTGAGGAATATGAAGTCCCACTGAAGAT 332
 Db 76 AGTGAATCATCCAAATCAAGCTATCCCTGTGAGGAATATGAAGTCCCACTGAAGAT 135
 QY 333 GGGTATATCTTCTCTTAAACAGGATTCCTCAGGCTAGTGCACCTTAAGACAGGT 392
 Db 136 GGGTATATCTTCTCTTAAACAGGATTCCTCAGGCTAGTGCACCTTAAGACAGGT 195
 QY 393 TCCAGGCTGTGGTGTACTGTCAGCATGGCTTAGTTGGAGGTGCTAGCAACTGGATTTC 452
 Db 196 TCCAGGCTGTGGTGTACTGTCAGCATGGCTTAGTTGGAGGTGCTAGCAACTGGATTTC 255
 QY 453 AACCTGCCCAACAATAGCTGGCTTCATTCTGGCAGATGCTGGTTTGAAGTGTGGATG 512
 Db 256 AACCTGCCCAACAATAGCTGGCTTCATTCTGGCAGATGCTGGTTTGAAGTGTGGATG 315

Db	1054	TTGACAGGAGGT	CAGGACTGGCTTTCAAAATCCAGAAAGACGTGAAAAATGCCTCTCTGAG	1111
Qy	1239	GTGACCAACCT	CATCTACCAATAAGATATATTCCTGAATGGGCTCACGTGGATTTTCATCTGG	1298
Db	1114	GTGACCAACCT	CATCTACCAATAAGATATATTCCTGAATGGGCTCACGTGGATTTTCATCTGG	1173
Qy	1299	GGTTTGGATGCT	CCTCCTCACCGTATGTTCAATGAATAATCATCATCTGATGCACGAGGAGGAG	1358
Db	1174	GGTTTGGATGCT	CCTCCTCACCGTATGTTCAATGAATAATCATCATCTGATGCACGAGGAGGAG	1233
Qy	1359	ACCAACCTTT	CCAGGACGGTGTGAGGCGGTATTCGTA	1397
Db	1234	ACCAACCTTT	CCAGGACGGTGTGAGGCGGTATTCGTA	1272
RESULT 9				
ABS55357				
XX	ABS55357	standard; DNA; 18554 BP.		
AC	ABS55357;			
XX	07-JAN-2003	(first entry)		
DT	DT			
XX	Genomic DNA	encoding human lipase protein.		
XX	Human; lipase;	cancer; colon; kidney; skin; brain; testis; chromosome 10;		
KW	Burkitt's lymphoma;	cytostatic; gene therapy; lipase-inhibitor; gene; ds;		
KW	single nucleotide polymorphism; SNP.			
XX	Homo sapiens.			
XX	XX			
PH	Key	Location/Qualifiers		
FT	CDS	2001..16554		
FT	FT	/*tag= a		
FT	FT	/product= "Human lipase protein"		
FT	FT	2001..2072		
FT	FT	/*tag= b		
FT	FT	/number= 1		
FT	FT	2073..4364		
FT	FT	/*tag= c		
FT	FT	/number= 1		
FT	FT	4365..4482		
FT	FT	/*tag= d		
FT	FT	/number= 2		
FT	FT	4483..9149		
FT	FT	/*tag= e		
FT	FT	/number= 2		
FT	FT	9150..9348		
FT	FT	/*tag= f		
FT	FT	/number= 3		
FT	FT	9349..10582		
FT	FT	/*tag= g		
FT	FT	/number= 3		
FT	FT	10583..10692		
FT	FT	/*tag= h		
FT	FT	/number= 4		
FT	FT	10693..11140		
FT	FT	/*tag= i		
FT	FT	/number= 4		
FT	FT	11141..11277		
FT	FT	/*tag= j		
FT	FT	/number= 5		
FT	FT	11278..11409		
FT	FT	/*tag= k		
FT	FT	/number= 5		
FT	FT	11410..11556		
FT	FT	/*tag= l		
FT	FT	/number= 6		
FT	FT	11557..12868		
FT	FT	/*tag= m		
FT	FT	/number= 6		
FT	FT	12869..12940		
FT	FT	/*tag= n		

FT intron /number= 7
 FT 12941..13877
 FT /tag= o
 FT /number= 7
 FT 13878..13949
 FT /tag= p
 FT /number= 8
 FT 13950..16284
 FT /tag= q
 FT /number= 8
 FT 16285..16551
 FT /tag= r
 FT /number= 9
 FT replace (1709,G)
 FT /tag= s
 FT /standard_name= "single nucleotide polymorphism (SNP)"
 FT replace (1816,A)
 FT /tag= t
 FT /standard_name= "single nucleotide polymorphism (SNP)"
 FT replace (1981,G)
 FT /tag= u
 FT /standard_name= "single nucleotide polymorphism (SNP)"
 FT replace (2174,A)
 FT /tag= v
 FT /standard_name= "single nucleotide polymorphism (SNP)"
 FT replace (2285,C)
 FT /tag= w
 FT /standard_name= "single nucleotide polymorphism (SNP)"
 FT replace (3069,C/G)
 FT /tag= x
 FT /standard_name= "single nucleotide polymorphism (SNP)"
 FT replace (3086,T)
 FT /tag= y
 FT /standard_name= "single nucleotide polymorphism (SNP)"
 FT replace (3104,G)
 FT /tag= z
 FT /standard_name= "single nucleotide polymorphism (SNP)"
 FT replace (3217,A)
 FT /tag= aa
 FT /standard_name= "single nucleotide polymorphism (SNP)"
 FT replace (3293,G)
 FT /tag= ab
 FT /standard_name= "single nucleotide polymorphism (SNP)"
 FT replace (3588,G)
 FT /tag= ae
 FT /standard_name= "single nucleotide polymorphism (SNP)"
 FT replace (4320,A)
 FT /tag= af
 FT /standard_name= "single nucleotide polymorphism (SNP)"
 FT replace (4960,T)
 FT /tag= ag
 FT /standard_name= "single nucleotide polymorphism (SNP)"
 FT replace (5722,A)
 FT /tag= ah
 FT /standard_name= "single nucleotide polymorphism (SNP)"
 FT replace (6018,G)
 FT /tag= ai
 FT /standard_name= "single nucleotide polymorphism (SNP)"
 FT replace (6562..6584,CA)
 FT /tag= aj
 FT /standard_name= "single nucleotide polymorphism (SNP)"
 FT replace (6573..6575AA)
 FT /tag= ak
 FT /standard_name= "single nucleotide polymorphism (SNP)"
 FT replace (6600,G)
 FT /tag= al
 FT /standard_name= "single nucleotide polymorphism (SNP)"

FT variation replace (10518,C)
 FT /tag= am
 FT /standard_name= "single nucleotide polymorphism (SNP)"
 FT replace (10857,C)
 FT /tag= an
 FT /standard_name= "single nucleotide polymorphism (SNP)"
 FT replace (1109,A)
 FT /tag= ao
 FT /standard_name= "single nucleotide polymorphism (SNP)"
 FT replace (13112,C)
 FT /tag= ap
 FT /standard_name= "single nucleotide polymorphism (SNP)"
 FT replace (15905,C)
 FT /tag= aq
 FT /standard_name= "single nucleotide polymorphism (SNP)"
 FT replace (15968,A)
 FT /tag= ar
 FT /standard_name= "single nucleotide polymorphism (SNP)"
 FT replace (16626,A)
 FT /tag= as
 FT /standard_name= "single nucleotide polymorphism (SNP)"
 FT XX
 PN W0200274975-A2.
 XX
 XX 26-SEP-2002.
 XX
 XX 18-MAR-2002; 2002WO-US08035.
 XX
 XX 20-MAR-2001; 2001US-0811825.
 XX
 XX (PEKE) PE CORP NY.
 XX
 XX Yan C, Di Francesco V, Beasley EM;
 XX
 XX WPI: 2002-750560/81.
 DR P-FSDB; ABG71000.
 XX
 XX New isolated human lipase peptides and encoding nucleic acids, useful
 PT for diagnosing and treating disorders mediated by human lipase
 PT proteins, e.g. cancer of the colon, kidneys, skin and testis, and
 PT Burkitt's lymphoma -
 PT
 XX Claim 4; Fig 3; 79pp; English.
 PS
 CC The present invention relates to a new lipase peptide. The methods and
 CC compositions of the present invention are useful for diagnosing and
 CC treating disorders mediated by the human lipase protein, such as cancer
 CC of the colon, kidneys, skin, brain and testis, and Burkitt's lymphoma.
 CC The present nucleic acid sequence represents the human lipase gene
 CC located on chromosome 10. This sequence encodes the human lipase protein
 CC of the invention.
 CC
 XX
 SQ Sequence 18554 BP; 5509 A; 3629 C; 3614 G; 5802 T; 0 other;
 Query Match 43.6%; Score 890.4; DB 24; Length 18554;
 Best Local Similarity 99.9%; Pred. No. 1.5e-217;
 Matches 891; Conservative 0; Mismatches 1; Indels 0; Gaps 0;
 QY 1126 AGCCAACTCCTGTAAGGTACAGATCAGAGATATACGGTCCCTACAGCAATGTGGACAG 1185
 DB 16283 AGCCAACTCCTGTAAGGTACAGATCAGAGATATACGGTCCCTACAGCAATGTGGACAG 16342
 QY 1186 GAGGTACAGACTGGCTTCCAAATCCAGAGAGCTGAAAATGCTCTCTGAGGTGACCA 1245
 DB 16343 GAGGTACAGACTGGCTTCCAAATCCAGAGAGCTGAAAATGCTCTCTGAGGTGACCA 16402
 QY 1246 ACCTCATCTACCATAGAAATATTCCTGTAATGGCTCAGCTGATTCATCTGGGGTTTGG 1305
 DB 16403 ACCTCATCTACCATAGAAATATTCCTGTAATGGCTCAGCTGATTCATCTGGGGTTTGG 16462
 QY 1306 ATGCTCCTCACCCTGATGTACAAATGAAATCCATCCATCTGATGAGGAGGACCAACC 1365
 DB 16463 ATGCTCCTCACCCTGATGTACAAATGAAATCCATCCATCTGATGAGGAGGACCAACC 16522

QY 633 TTTATTTTTCAGAAACGGCCAGGAAAGATCTATTATGCTGCTATTACAGGGCACC 592
 Db 433 TTTATTTTTCAGAAACGGCCAGGAAAGATCTATTATGCTGCTATTACAGGGCACC 492
 QY 693 ACCATGGGCTTTATTTGCAATTTCCACCATGCCAGAGCTGGCTCAGAAAATCAAAATGTAT 752
 Db 493 ACCATGGGCTTTATTTGCAATTTCCACCATGCCAGAGCTGGCTCAGAAAATCAAAATGTAT 552
 QY 753 TTTGCTTTAGACCCCATAGCACTGTTAGCATGCAAAAAGCCCGGACCAAAATTTTG 812
 Db 553 TTTGCTTTAGACCCCATAGCACTGTTAGCATGCAAAAAGCCCGGACCAAAATTTTG 612
 QY 813 TTTGCTGCGAGATGATGATCAAGGGATTTTGGCAAAAAGAAATTTCTGTATCAGACC 872
 Db 613 TTTGCTGCGAGATGATGATCAAGGGATTTTGGCAAAAAGAAATTTCTGTATCAGACC 672
 QY 873 AGATTTCTCAGACAACTTGATTTTACCTTTGCGGAGTGTATTTCTGATCAGATTTGT 932
 Db 673 AGATTTCTCAGACAACTTGATTTTACCTTTGCGGAGTGTATTTCTGATCAGATTTGT 732
 QY 933 AGTAATATCATGTTTACTTTCTGGGTGGATTCACACCAACAATATGAACAT----- 982
 Db 733 AGTAATATCATGTTTACTTTCTGGGTGGATTCACACCAACAATATGAACATGAATCTCAT 792
 QY 983 -----GAGCCGAGCAAGTGTATATGCTGCCACACTTTCTGCGAACATCTGTG 1031
 Db 793 GGTGTTGTAAGAGCCGAGCAAGTGTATATGCTGCCACACTTTCTGCGAACATCTGTG 852
 QY 1032 CAAAATATTCTACACTGGAGCCAG 1055
 Db 853 CAAAATATTCTACACTGGAGCCAG 876

RESULT 11.

ID ABN97383 standard; DNA; 2481 BP.
 XX AC ABN97383;
 XX DT 13-AUG-2002 (first entry)
 XX DE Gene #3881 used to diagnose liver cancer.
 XX KW Gene; liver cancer; ds; hepatocellular carcinoma; hepatotropic;
 KW metastatic liver tumor; cytotoxic; expression profile; disease state;
 KW disease progression; drug toxicity; drug efficacy; drug metabolism.
 XX OS Homo sapiens.
 XX XN WO200229103-A2.
 XX PD 11-APR-2002.
 XX PF 02-OCT-2001; 2001WO-US30589.
 XX PR 02-OCT-2000; 2000US-237054P.
 XX PA (GENE-) GENE LOGIC INC.
 XX PI Horne D, Alvares C, Perea-Da-Silva S, Vockley JG;
 XX DR WPI; 2002-426119/45.
 XX PT Diagnosing and detecting the progression of liver cancer,
 PT hepatocellular carcinoma or metastatic liver tumor in a patient,
 PT involves detecting the level of expression of two or more genes in a
 PT liver tissue sample -
 XX PS Claim 1; SEQ ID NO 3881; 298pp; English.
 CC The invention relates to a novel method for diagnosing and detecting the
 CC progression of liver cancer, hepatocellular carcinoma or metastatic liver
 CC tumour in a patient, and differentiating metastatic liver cancer from

CC hepatocellular carcinoma in a patient, involving detecting the level of
 CC expression of two or more genes represented in ABN97383-ABN97455 in a
 CC tissue sample. The method of the invention has hepatotropic, and
 CC cytotoxic activity. The method is useful for diagnosing and detecting
 CC the progression of liver cancer, hepatocellular carcinoma and metastatic
 CC liver carcinoma in a patient. The method is useful for identifying
 CC expression profiles which serve as useful diagnostic markers as well as
 CC markers that can be used to monitor disease states, disease progression,
 CC drug toxicity, drug efficacy and drug metabolism.
 CC Note: The sequence data for this patent did not form part of the printed
 CC specification, but was obtained in electronic format directly from WIPO
 CC at ftp.wipo.int/pub/published_pat_sequences.
 XX SQ Sequence 2481 BP; 692 A; 477 C; 513 G; 799 T; 0 other;
 Query Match 25.8%; Score 527.8; DB 24; Length 2481;
 Best Local Similarity 66.6%; Pred. No. 6.2e-125;
 Matches 751; Conservative 0; Mismatches 376; Indels 0; Gaps 0;
 QY 240 AAAGCTGTGGACCCAGAGGATTCATGAATATTAGTGAATATCAATCAAGGCTAT 299
 Db 112 ACAGCTGTGGATCTCTGAAACAAACATCAATGTGAGTGAATATCTCTTCTGGGATTC 171
 QY 300 CCCTGTGAGGAATATGAGTGCACCTGCAAGATGGGTATATCTTCTGTAAACAGATT 359
 Db 172 CCTAGTGAAGGAATACCTAGTTGAGACAGAGATGATATATTCTGCTTAAACGATTT 231
 QY 360 CCTGAGGCTTGTGCAACCTTGAAGACAGAGTTCAGGCTGTGGTGTACTGACGAT 419
 Db 232 CCTCATGGGAGGAGAAACCATTTCTGACAAAGGTCCTCAACAGTTGTCTTCTGCAACAT 291
 QY 420 GGCTGTGAGGATGTGAGTGCACCTGCAAGATGGGTATATCTTCTGTAAACAGATT 479
 Db 292 GGCTGTGAGGATGTGAGTGCACCTGCAAGATGGGTATATCTTCTGTAAACAGATT 351
 QY 480 ATTTCTGAGATGCTGTTTGTGATGGGAAACAGAGGAGGAAACGCTGTCT 539
 Db 352 ATTTCTGAGTGTGTTTGTGATGGGAAACAGAGGAGGAAACGCTGTCT 411
 QY 540 CGAAACACAGACACTCTCATAGACAGAGTGTCTGGGCTTTCAGTTATGATGAG 599
 Db 412 CGAAACACAGACACTCTCATAGACAGAGTGTCTGGGCTTTCAGTTATGATGAG 471
 QY 600 ATGCTAGTGTGACCTTCTGCAAGTGAATAACTTTATTTTGCAGAAAACGGCCAGAA 659
 Db 472 ATGCAAAATATGACCTTACCAAGTTCCTTCAATTAATTTCTGAAATAAACTGGCCAAAGAA 531
 QY 660 AAGATCTATTATGCTGCTATTACAGGGACCAACAGGCTTTATTTGATTTTCCACC 719
 Db 532 CAAGTGTATTATGCTGCTATTACAGGGACCAACAGGCTTTATTTGATTTTCCAG 591
 QY 720 ATGCCAGAGCTGGCTCAGAAAATCAAAATGTATTTTCTTTAGCACCCTAGCCATGTT 779
 Db 592 ATCCCTGAGCTGGCTTAAAGGATTTAAATGTTTTTGGCTGCTGCTGCTGCTGCTGCT 651
 QY 780 AAGCATGCAAAAGCCCGGAGCAAAATTTTGTGCTGCCAGATATGATGATCAAGGA 839
 Db 652 GCTTCTGCTAGTCCCTATGCGCAAAATAGGACGATTAACAGATCACTCAATTAAGGAC 711
 QY 840 TTGTTTGGCAAAAAGAAATTTCTGTATCAGACCAAGATTTCTCAGACCAACTTTTATTAC 899
 Db 712 TTATTTGGAGCAAGAAATTTCTTCCACAGAGTGGTTTTTGAAGTGGCTGGGTACCCAC 771
 QY 900 CTTTGTGGCAGGATGATTTCTGATCAGATTTGTAGTAAATATATCATTTCTCTGGTGA 959
 Db 772 GTTTCACACTCATCTACTAGGAGCTCTGTGAAAATCTCTGTTTTTCTCTGTGGA 831
 QY 960 TTCAACACCAACAATATGAACATGAGCGAGCAAGTGTATATGCTGCCACACTCTTGTCT 1019
 Db 832 TTATATGAGAAATTTTAAATATGCTAGATGATGATATATCAACACATTTCTCTGCT 891
 QY 1020 GGAAATCTGTGCAAAATATTTCTACCTGAGGACGAGGAAATTTCTGTGGAACCTCCGG 1079

Db 892 GGAACCTCTGTGCAAAAACATGTTACATGAGCCAGGCTGTTAAATTCAAAAGTTTCAA 951
 QY 1080 GCATTGAGCTGGGGAGTGAGACCAAAATCTGGAAAATCAATCAACCAACTCCTGTA 1139
 Db 952 GCCTTGAGCTGGGAGCGAGTCCAGAAATATTTTCAATCAACCAAGTTATCCTCCC 1011
 QY 1140 AGGTACAGAGTCAGAGATATGACGGTCCCTACAGCAATGTGACAGGAGTCAAGACTGG 1199
 Db 1012 ACATACAATGTGAAGGACATGCTTGTGCGACTGCGAGTCTGAGCGGGGGTCAAGACTGG 1071
 QY 1200 CTTTCAAAATCCAGAAGAGCTGAAAATGCTGCTCTGAGGTGACCAACTCATCTACCAT 1259
 Db 1072 CTTGAGAGTCTACAGAGCTCAATATCTTACTGATCTCAGATCACCACCTTGGTGTCCAT 1131
 QY 1260 AAGAAATATTCGAATGGGCTCACGTGGATTTTCATCTCGGGTTTGGATGCTCCTCACCGT 1319
 Db 1132 GAGAGCAATCCGGAATGGGAGCATCTTGACTTCAATTTGGGGCTGGATGCCCTTGGAGG 1191
 QY 1320 ATGTACAATGAATCATCTCATCTGATGACGAGGAGGACCAACCT 1366
 Db 1192 CTTTATAATAAATTAATTAATCTAATGAGGAATATCAGTGAAGCT 1238

RESULT 12

AAH02906

ID AAH02906 standard; DNA; 2493 BP.

XX AC AAH02906;

XX 15-JUN-2001 (first entry)

XX Human shear stress-response coding sequence SEQ ID NO: 65.

XX Human; shear stress-response protein; vascular disease;

XX arteriosclerosis; ds.

XX Homo sapiens.

XX WO200125427-A1.

XX 12-APR-2001.

XX 02-OCT-2000; 2000WO-JP06840.

XX 01-OCT-1999; 99JP-0280976.

XX (KYOW) KYOWA HAKKO KOGYO KK.

XX (NOJI/) NOJIMA H.

XX Nojima H, Yoshisue H, Obayashi M, Ota T, Kawabata A, Sakurada K;

PI Kuga T, Sekine S, Nakamura Y, Sugano S;

XX WPI; 2001-266308/27.

XX P-PSDB; AAB90783.

XX DNA sequences, proteins encoded by them and antibodies against them

XX useful in diagnosis and treatment of vascular disease caused by

XX arteriosclerosis -

XX Claim 20; Page 398-402; 678pp; Japanese.

XX The present invention provides the protein and coding sequences of a

XX number of human shear stress response proteins. These are useful in the

XX diagnosis, treatment and screening of vascular diseases caused by

XX arteriosclerosis, including heart failure, post-PTCA restenosis and

XX hypertension.

XX Sequence 2493 BP; 692 A; 484 C; 516 G; 801 T; 0 other;

XX Query Match

XX Best Local Similarity 25.7%; Score 526.2; DB 22; Length 2493;

XX Matches 751; Conservative 0; Mismatches 376; Indels 0; Gaps 0;

QY 240 AAAGCTGTGACCCAGCAATTCATGAATATTAGTGAATCATCCAAATCAAGGCTAT 299
 Db 119 ACAGCTGTGGATCCTGAACAAACATGAATGTGAGTGAATATATCTTCTTACTGGGATTC 178
 QY 300 CCTGTGAGGAATATGAGTCGCACTGAAGATGGGTATATCTTCTTCTGTTAACAGATT 359
 Db 179 CCTAGTGAAGGAATACCTAGTGTGAGACAGAAATGGAATATATCTGTGCTTAAACCGAAT 238
 QY 360 CTTCTGAGGCTTACTGTCAAACCTAAGAAGACAGAGTTCAGGCTCTGTGGTGTATCTGACGAT 419
 Db 239 CCTCATGGGAGGAAGAACCATCTCGACAAGGTCCTCAACCAAGTTGTCTTCTCTGCAACAT 298
 QY 420 GGCTAGTGTGAGGTGCTAGCAACTCGATTTCCAACTGCCCCAAATAGAGCTGGGCTTC 479
 Db 299 GGCTTGTGTCAGATTTCTAGTAACTGGGTCAAAACCTTTGCCAACAGAGCTGGGCTTC 358
 QY 480 ATTCTGCGCATGCTGCTTTTTCAGCTGTGATGGGAAACAGAGGGGAAACGCTGCTCT 539
 Db 359 ATTCTTGTCTGATGCTGCTTTTTCAGCTGTGATGGGAAACAGAGAGGAATACCTGCTCT 418
 QY 540 CGAAACACAAAGACATCTCTCCATAGACCAAGATGAGTTCCTGGGCTTTCAGTTATGATGAG 599
 Db 419 CGGAAACATAGAAGACATCTCTCAGTTTCTCAGGATGAATTCGGGCTTTCAGTTATGATGAG 478
 QY 600 ATGCTAGGTTTTCACCTTCTCTGAGTGATAAATCTTTTATTTTGAGAAAACGGGCCAGGAA 659
 Db 479 ATGGCAAAATATGACCTTACAGCTTCATTAATCTTCAATCTGAAATAAAATCGGCCAAGAA 538
 QY 660 AAGATCTATTATGTCGGCTATTACAGGGCACCAAGGCTTTTATGATGATGATGATGATGATGAT 719
 Db 539 CAAGTGTATTATGTCGGCTATTCTCAAGGCACCACTATAGGTTTATATAGCATTTTTCACAG 598
 QY 720 ATGCGAGCTGGCTCAGAAAATCAAAATGTAATTTTGTCTTTAGCACCCATAGCACTGTT 779
 Db 599 ATCCCTGAGCTGGCTAAAAGGATTAATATGTTTTTGGCCCTGGGCTCTCTGGGCTCCGTC 558
 QY 780 AAGCATGCAAAAAGCCCGGGACCAAATTTTGTCTGCCAGATATGATGATGATGATGATGATGATGAT 839
 Db 659 GCCTTCTGTACTAGCTTACCTGAGGCTCTGTGGAAATCTCTGTTTTCTTCTGTGTGGA 718
 QY 840 TTGTTTGGCAAAAAGAAATTTCTGTATCAGACCAAGATTTCTCAGACAACTTGTATTATAC 899
 Db 719 TTATTTGGAGACAAAGAAATTTCTTCCCGAGAGTGCCTTTTGAAGTGGCTGGGTACCCAC 778
 QY 900 CTTTGTGGCCAGGTGATTTCTTGATCAGATTTGTAGTAATATCATGTTACTTCTTGGGTGGA 959
 Db 779 GTTTGCACTCATGTCATCTGAGGAGCTCTGTGGAAATCTCTGTTTTCTTCTGTGTGGA 838
 QY 960 TTCAACACCAATATGAACATGAGCCGAGAGCAAGTGTATATGCTGCCACCACTCTTGCT 1019
 Db 839 TTTAATGAGAGAAATTTAAATATGTTCTAGAGTGGATGTATATACAAACATCTCTCTGCT 998
 QY 1020 GGAACATCTGTGCAAAATATTTACACTGAGCGGAGGAGTGAATTTCTGGTGAATCTCCGG 1079
 Db 899 GGAATCTGTGCAAAACATGTTTACACTGAGCGGAGGCTGTTAAATTTCCAAAGTTTCAA 958
 QY 1080 GCATTGAGCTGGGGAGTGAGACCAAAATCTGGAAAATGCAATCAGCAACTCTCTGTA 1139
 Db 959 GCCTTTGACTGGGGAAGCAGTGCCTAAGAAATTTTTCATTACCAACAGAGTTATCTCTCC 1018
 QY 1140 AGGTACAGAGTCAGAGATATGACGGTCCCTACAGCAATGTGACAGAGGCTCAGGACTGG 1199
 Db 1019 ACATACATGTGAGAGCATGCTTGTGGCAGCTGCGAGTGTGAGCGGGGGTCAAGACTGG 1078
 QY 1200 CTTTCAAAATCCAGAGACGTTGAAAATGCTGCTCTCTGAGGTGACCAACCTCACTACCT 1259
 Db 1079 CTTGCAATGCTTACGACGTCATATCTTACTGACTCAGATCACCACCTTGGTGTGCTCAT 1138
 QY 1260 AAGATATTTCTGAATGGGCTCAGGTGATTTTCACTCTGGGTTTGGATGCTCTCTCAGCT 1319
 Db 1139 GAGAGCAATCCGGAATGGGAGCATCTTGACTTCAATTTGGGGCTGGATGCCCTTGGAGG 1198
 QY 1320 ATGTACAATGAATCATCTCATCTGATGACGAGGAGGAGACCAACCT 1366

PT mat_peptide 104..1243
 FT /*tag= b
 XX W08601532-A.
 XX 13-MAR-1986.
 XX 15-AUG-1985; 85WO-GB00364.
 XX 21-AUG-1984; 84GB-0021210.
 XX 15-AUG-1985; 85WO-GB00364.
 XX 01-JAN-1986; 86GB-0008957.
 XX (CLLT) CELLTECH LTD.
 PA (LOWE/) LOWE P A.
 XX Lowe PA;
 XX WPI; 1986-081634/12.
 DR P-PSDB; AAP60658.
 XX
 PT New gastric lipase protein, esp. of human origin - for treating
 PT lipase deficiency, and DNA sequences coding for it
 XX
 PS Disclosure; Fig 3; 39pp; English.
 XX
 CC The inventors claim a pregastric lipase protein and a gene encoding
 CC it. Gastric lipase protein is useful for oral administration to
 CC treat lipase deficiency, e.g. cystic fibrosis or pancreatitis.
 CC (Updated on 25-MAR-2003 to correct PA field.)
 XX
 XX Sequence 1367 BP; 397 A; 284 C; 269 G; 417 T; 0 other;
 SQ
 Query Match 23.0%; Score 470; DB 7; Length 1367;
 Best Local Similarity 64.2%; Pred. No. 3.2e-110;
 Matches 707; Conservative 0; Mismatches 395; Indels 0; Gaps 0;
 244 CTGTGGACCCAGAGCATTATGATGATATAGTGAATCATCAACATCAAGGTATCCCT 303
 123 CTGGAAGCCCTGAAGTCACTATGACATATAGTGAATGATGATGATGATGATGATGAT 182
 304 GTGAGGAATATGAAGTCCCACTCAAGATGGGTATATCCCTTCTGTAAACAGGATTCCTC 363
 183 ATGAAGATATGAAGTGTGACTGAAGATGTTTATATCTTGAAGTCAATGATATCCCT 242
 364 GAGCCTAGTGCACCTAAGAGACAGGTTCAGGCTGCTGCTGCTGCTGCTGCTGCTGCTG 423
 243 ATGGGAAGAAAAATTCAGGGAATACAGGCCAGAGACCTGTTGTTGTTGTTGTTGTTGTT 302
 424 TAGTTGAGGTGCTAGCAACTGGATTTCCAACTGCCCAACAATAGCTGGGCTTCATTC 483
 303 TGCTTGATCAGCCACAACTGGATTTCCAACTGCCCAACAATAGCTGGGCTTCATTC 362
 484 TGGCAGATGCTGGTTTGAAGTGGATGGGGAACAGCAGGGGAACAGCCTGCTGCTGCTG 543
 363 TGGCAGATGCTGGTTTGAAGTGGATGGGGAACAGCAGGGGAACAGCCTGCTGCTGCTG 422
 544 AACACACACACTCTCCATACCAAGATGATGATGATGATGATGATGATGATGATGATGATG 603
 423 GAACCTGCTATATCACCAGATTCAGTTGATTTCTGGGCTTTCAGGCTTGTGATGATG 482
 604 CTAGTTTGCATCTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 663
 483 CTAAATATGACCTTCCAGCCCAATTCGACTTCAATGTAAGAAAGAAATCTGGACAGAG 542
 664 TCTATTATGCTGCTATTCAGAGGACCACTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 723
 543 TACACTATGTTGGCCATTCAGAGGACCACTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 602
 724 CAGAGCTGGCTCAGAAAAATCAAAATGATTTTGTCTTAGCACCCTAGCCACTGTTAAGC 783
 603 CCAGCTGGCTAAAAGATCAAAACCTTCTATGCTCTAGCTCTGCTGCTGCTGCTGCTGCTG 662

QY 784 ATGCAAAAGCCCGGACCAAAATTTTGTGTCGACAGATATGATGATCAAGGGATTGT 843
 DB 663 ATACAAAAGCCTTATAAACAACTTAGATTGTTCTCTCAATCCCTCTTCAAGTTTATAT 722
 QY 844 TTGCAAAAAAGAAATTTCTGTATCAGACCAGATTCTCTCAGACAACTTGTATTATTACCTTT 903
 DB 723 TTGGTGAACAAATATTTCTAGCCACACAACTTCTTGTATCAATTTCTTGTCTACTGAAGTGT 782
 QY 904 GTGGCCAGGTGATTTCTTGTATCAGATTTGTAGTATATATCATGTTACTTCTGGGTGGATTC 963
 DB 783 GCTCCCGTGAGTGTGAATCTCCTTTGACGCAATGCTTATTTAATTTTGGATTG 842
 QY 964 ACACCAACAATATGAACATGAGCGAGCAAGTGTATGCTGCCACACACTCTTCTGCTGGAA 1023
 DB 843 ACAGTAAGAACTTTAAACACAGTCTGCTTGGATGTATCTATCACATAATCCACGAGAA 902
 QY 1024 CATCTGTCAAAATATTTCTACACTGGAGCCAGGAGTGAATTTCTGGTGAACCTCCGGCAT 1083
 DB 903 CTTCTGTTCAAAACATGTTCCATTTGGACCCAGGCTGTTAAGTCTGGGAATTCAGACTT 962
 QY 1084 TTGACTGGGGAGTGAACCAAAATCTGGAATAATCAATCAAGCCAACTCTCTGTAAAGT 1143
 DB 963 ATGACTGGGGAAGCCAGTTTCAGAAATAGGATGCACTATGATCAGTCCCAACCTCCCTACT 1022
 QY 1144 ACAGAGTCAGAGATATGACGTCCTCTACAGCAATGTGGACAGGAGTCAAGGCTGCTT 1203
 DB 1023 ACAATGTGACAGCCATGAATGTACCAATTCAGTGTGGAAACGGTGGGAGAGACTGTTGG 1082
 QY 1204 CAAATCCAGAGAGAGTGAATAATGCTGCTCTCTGAGGTGACCAACCTCATCTACCATAGA 1263
 DB 1083 CTGACCCCAAGATGTTGGGCTTTTGTCTCCAAACTCCCAATCTTATTTTACCAAGG 1142
 QY 1264 ATATTCTGAAATGGGCTCAGTGGATTTCACTCTGGGCTTGGATGCTCTCTCAGCGTATGT 1323
 DB 1143 AGATTCTCTTTTACAATCACTTGGACTTATCTGGGCAATGGATGCCCCCTCAAGAGTTT 1202
 QY 1324 ACAATGAATCATCCATCTGAT 1345
 DB 1203 ACAATGACATTTTCTATGAT 1224

Search completed: February 19, 2004, 05:05:57
 Job time : 572.905 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2004 Compugen Ltd.

OM nucleic - nucleic search, using sw model

Run on: February 19, 2004, 04:48:12 ; Search time 4470.52 Seconds
(without alignments)
11112.438 Million cell updates/sec

Title: US-10-042-431-45

Perfect score: 2044

Sequence: 1 gtgacccacgcgtccgggg.....aaaaaaaaaagggcgccgc 2044

Scoring table:

IDENTITY_NUC
Gapop 10.0 , Gapext 1.0

Searched: 22781392 seqs, 12152238056 residues

Total number of hits satisfying chosen parameters: 45562784

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database :

E8T:*

1: em_estba:*

2: em_esthum:*

3: em_estin:*

4: em_estmu:*

5: em_estov:*

6: em_estpl:*

7: em_estro:*

8: em_htc:*

9: gb_est1:*

10: gb_est2:*

11: gb_htc:*

12: gb_est3:*

13: gb_est4:*

14: gb_est5:*

15: em_estfun:*

16: em_estom:*

17: em_gss_hum:*

18: em_gss_inv:*

19: em_gss_pln:*

20: em_gss_vrt:*

21: em_gss_fun:*

22: em_gss_mam:*

23: em_gss_mus:*

24: em_gss_pro:*

25: em_gss_rtd:*

26: em_gss_pbg:*

27: em_gss_vri:*

28: gb_gss1:*

29: gb_gss2:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	1005.2	49.2	2050	11 AK037091	AK037091 Mus muscu
2	1005.2	49.2	2678	11 AK085719	AK085719 Mus muscu
3	1003.6	49.1	2555	11 AK037214	AK037214 Mus muscu
4	628.2	30.7	2527	11 AK019504	AK019504 Mus muscu

C	5	592	29.0	629	10	BG743024
C	6	543	26.6	556	13	BX103621
	7	505	24.7	954	10	BF135102
	8	497.6	24.3	1079	10	BF141567
	9	492	24.1	669	10	BF143546
	10	475.4	23.3	1338	11	AK088659
	11	461.8	22.6	1350	11	AK010093
	12	460.2	22.5	1347	11	AK010116
	13	460.2	22.5	1349	11	AK009300
	14	460.2	22.5	1350	11	AK010139
	15	458.6	22.4	1340	11	AK010124
	16	458.6	22.4	1343	11	AK009479
	17	458.6	22.4	1345	11	AK009413
	18	458.6	22.4	1345	11	AK009473
	19	458.6	22.4	1346	11	AK009359
	20	458.6	22.4	1346	11	AK009523
	21	458.6	22.4	1346	11	AK009729
	22	458.6	22.4	1348	11	AK010058
	23	458.6	22.4	1349	11	AK009474
	24	458.6	22.4	1349	11	AK009546
	25	458.6	22.4	1349	11	AK009571
	26	458.6	22.4	1349	11	AK009573
	27	458.6	22.4	1349	11	AK010148
	28	458.6	22.4	1350	11	AK009459
	29	458.6	22.4	1350	11	AK009525
	30	458.6	22.4	1350	11	AK009773
	31	458.6	22.4	1350	11	AK010061
	32	458.6	22.4	1350	11	AK010125
	33	458.6	22.4	1350	11	AK010203
	34	458.6	22.4	1351	11	AK009428
	35	458.6	22.4	1352	11	AK009431
	36	458.6	22.4	1356	11	AK010019
	37	458.6	22.4	1356	11	AK010035
	38	457.6	22.4	1346	11	AK009437
	39	457.6	22.4	1347	11	AK010103
	40	457.6	22.4	1349	11	AK010026
	41	457	22.4	1346	11	AK009560
	42	457	22.4	1347	11	AK010236
	43	457	22.4	1348	11	AK010231
	44	457	22.4	1349	11	AK009875
	45	457	22.4	1350	11	AK009537

ALIGNMENTS

RESULT 1
LOCUS AK037091
DEFINITION Mus musculus adult female vagina cDNA, RIKEN full-length enriched library, clone:993011004 product:BA30415.1 (NOVEL LIPASE) (FRAGMENT) homolog [Homo sapiens], full insert sequence.
ACCESSION AK037091
VERSION AK037091.1 GI:26331937
KEYWORDS HTC; CAP trapper.
SOURCE Mus musculus (house mouse)
ORGANISM Mus musculus
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
AK037091 2050 bp mRNA linear HTC 05-DEC-2002

REFERENCE

AUTHORS

TITLE

JOURNAL

MEDLINE

PUBMED

REFERENCE

AUTHORS

TITLE

JOURNAL

MEDLINE

PUBMED

REFERENCE

AUTHORS

TITLE

JOURNAL

MEDLINE

PUBMED

Carninci, P. and Hayashizaki, Y.
High-efficiency full-length cDNA cloning
Meth. Enzymol. 303, 19-44 (1999)
93273253
10349636
Carninci, P., Shibata, Y., Hayatsu, N., Sugahara, Y., Shibata, K., Ichikawa, M., Kono, H., Okazaki, Y., Muramatsu, M. and Hayashizaki, Y.
Normalization and subtraction of cap-trapper-selected cDNAs to prepare full-length cDNA libraries for rapid discovery of new genes
Genome Res. 10 (10), 1617-1630 (2000)
20499374
11042159

REFERENCE
AUTHORS

3 Shibata, K., Itoh, M., Aizawa, K., Nagaoka, S., Sasaki, N., Carninci, P., Konno, H., Akiyama, J., Nishi, K., Katsunuma, T., Ieshiro, H., Itoh, M., Sumi, N., Ishii, Y., Nakamura, S., Hazama, M., Nishino, T., Harada, A., Yamamoto, R., Matsumoto, H., Sakaguchi, S., Ikegami, T., Kashiwagi, K., Fujitake, S., Inoue, K., Togawa, Y., Izawa, M., Ohara, E., Watahiki, M., Yoneda, Y., Ishikawa, T., Ozawa, K., Tanaka, T., Matsura, S., Kawai, J., Okazaki, Y., Muramatsu, M., Inoue, Y., Kira, A. and Hayashizaki, Y.
RIKEN integrated sequence analysis (RISA) system-384-format
sequencing pipeline with 384 multicapillary sequencer
Genome Res. 10 (11), 1757-1771 (2000)
20530913
11076861

TITLE

JOURNAL

MEDLINE

PUBMED

REFERENCE
AUTHORS

4 Kawai, J., Shinagawa, A., Shibata, K., Yoshino, M., Itoh, M., Ishii, Y., Aizawa, K., Hara, A., Fukunishi, Y., Konno, H., Adachi, J., Fukuda, S., Arakawa, T., Izawa, M., Nishi, K., Kiyosawa, H., Kondo, S., Yananaka, I., Saito, T., Okazaki, Y., Gojobori, T., Bono, H., Kasukawa, T., Saito, R., Kadota, K., Matsuda, H., Ashburner, M., Batalov, S., Casavant, T., Fleischmann, W., Gaasterland, T., Gissi, C., King, B., Kochiwa, H., Kuehl, P., Lewis, S., Matsuo, Y., Nikaide, I., Pesole, G., Quackenbush, J., Schram, L. M., Staubli, F., Suzuki, R., Tomita, M., Wagner, L., Washio, T., Sakai, K., Okido, T., Furuno, M., Aono, H., Baldarelli, R., Barsh, G., Blake, J., Boffelli, D., Bojunga, N., Caldarini, P., de Bona, M. F., Brownstein, M. J., Bult, C., Fletcher, C., Fujita, M., Gariboldi, M., Gustincich, S., Hill, D., Hofmann, M., Hume, D. A., Kamiya, M., Lee, N. H., Lyons, P., Marchionni, L., Mashima, J., Mazzarelli, J., Mombaerts, P., Nordone, P., Ring, B., Ringwald, M., Rodriguez, I., Sakamoto, N., Sasaki, H., Sato, K., Schonbach, C., Seva, T., Shibata, Y., Storch, K. F., Suzuki, H., Toyooka, K., Wang, K. H., Weitz, C., Whittaker, C., Wilming, L., Wyshaw-Boris, A., Yoshida, K., Hasegawa, Y., Kawaji, H., Kohtsuki, S. and Hayashizaki, Y.
Functional annotation of a full-length mouse cDNA collection
Nature 409 (6821), 685-690 (2001)
21085660
11217851

TITLE

JOURNAL

MEDLINE

PUBMED

REFERENCE
AUTHORS

5 The PANTOM Consortium and the RIKEN Genome Exploration Research Group Phase I & II Team.
Analysis of the mouse transcriptome based on functional annotation of 60,770 full-length cDNAs
Nature 420, 563-573 (2002)
6 (bases 1 to 2050)
Adachi, J., Aizawa, K., Akimura, T., Arakawa, T., Bono, H., Carninci, P., Fukuda, S., Furuno, M., Hanagaki, T., Hara, A., Hashizume, W., Hayashida, K., Hayatsu, N., Hiramoto, K., Hiraoka, T., Hirozane, T., Hori, P., Imotani, K., Ishii, Y., Itoh, M., Kagawa, I., Kasukawa, T., Katoh, H., Kawai, J., Kojima, Y., Kondo, S., Konno, H., Kouda, M., Koya, S., Kurihara, C., Matsuyama, T., Miyazaki, A., Murata, M., Nakamura, M., Nishi, K., Nomura, K., Numazaki, R., Ohno, M., Ohsato, N., Okazaki, Y., Saito, R., Saitoh, H., Sakai, C., Sakai, K., Sakazume, N., Sano, H., Sasaki, D., Shibata, K., Shinagawa, A., Shiraki, T., Sogabe, Y., Tagami, M., Tagawa, A., Takahashi, F., Takaku-Akahira, S., Takeda, Y., Tanaka, T., Tomaru, A., Toya, T., Yasunishi, A., Muramatsu, M. and Hayashizaki, Y.
Direct Submission
Submitted (16-JUL-2001) Yoshihide Hayashizaki, The Institute of Physical and Chemical Research (RIKEN), Laboratory for Genome Exploration Research Group, RIKEN Genomic Sciences Center (GSC), RIKEN Yokohama Institute, 1-7-22 Suehiro-cho, Tsurumi-Ku, Yokohama, Kanagawa 230-0045, Japan [E-mail: genome-res@sc.riken.go.jp, URL: http://genome.gsc.riken.go.jp/, Tel: 81-45-503-9222, Fax: 81-45-503-9216]

TITLE

JOURNAL

COMMENT

cDNA library was prepared and sequenced in Mouse Genome Encyclopedia Project of Genome Exploration Research Group in Riken Genomic Sciences Center and Genome Science Laboratory in RIKEN, Division of Experimental Animal Research in Riken contributed to prepare mouse tissues.
Please visit our web site for further details.
URL: http://genome.gsc.riken.go.jp/
URL: http://fantom.gsc.riken.go.jp/
Location/Qualifiers

FEATURES

source

1. .2050
/organism="Mus musculus"
/mol_type="mRNA"
/strain="CS7BL/6J"
/db_xref="FANTOM_DB:930111004"
/db_xref="taxon:10090"
/clone="930111004"
/sex="female"
/tissue_type="vagina"
/clone_lib="RIKEN full-length enriched mouse cDNA library"
/dev_stage="adult"
144..1412
/notes="unnamed protein product; BA30415.1 (NOVEL LIPASE) (FRAGMENT) homolog (Homo sapiens) (SPTR) Q96LG2, evidence: FASTY, 88.5% ID, 100% length, match=1119)
putative"
/codon_start=1
/protein_id="BAC29699.1"
/db_xref="GI:26331938"
/translations="MSEILSRVTVSHRVEIWLIIIVALLQNVNSHLPKAAADPE AFNVSEIIKHGYPSEEEVATEDGIIISVNRIPKQTRLKKEGSRFVVLQHLGLG DSNWISINPNNSLGFILADGPDVWMSNRGTSRKHKLSIDODEFWAFSDYEMA RFDLPVINFILQTKQKYYVYSGQTMGFIAFTWPELAHKIMYFALAPIATV KYARSPGTEKLLPDMWIKVLFGRQFLYTFQFRLFYLCQIMLDGICNSIIILL GGFTNNMNSRANVVAHTPACTSVONILHWSQVNSGELRAFDMGSETKNOEKNO PTIRYKVRDMVPTAMWTGGQDLSNPDVDKTLSEVTLNLIYHKNIPFWAHVDFIWG LDAPQRYNEIILHMKQEPNLPQGTCEVLL"

CDS

polyA_signal
2027..2032
/note="putative"
polyA_site
2050
/note="putative"
BASE COUNT 598 a 415 c 461 g 576 t
ORIGIN

Query Match 49.2%; Score 1005.2; DB 11; Length 2050;
Best Local Similarity 84.0%; Pred. No. 1.5e-152;
Matches 1173; Conservative 0; Mismatches 213; Indels 10; Gaps 3;
QY 48 TTTTAAACCCACAAATCTTCTTACCTTTAGAAATGTTGTTTACATTTGCGAGGAA-AAAA 106
Db 65 TCTTTCATCTCTGTTCTTCTTCTGACTAGATGATGTTGTTACATCTTAGCAACAAG 124
QY 107 TAATGCGAGATGTTGGACCAATGTTGGAAACCTTGTCAAGACAGTGGATGTTCTCAGACAG 166
Db 125 CAATGCGCGATGTTGGACCAATGTCAGAAATCTTTTCAAGAGTGTGGACTGTTTGGCAG 184
QY 167 AATGGAATGCGCTTCTGATTTCTGTTGGGCTATGTTCCAGAGAAATGTAATTCAGT 226
Db 185 AGTGGAGATAGGTTCTGATTTCTGTTAGCGTATTTACTCCAGAAATGTAATTCAGTGGG 244
QY 227 ACATATGCCAACTAAAGCTGTGGACCCAGAACGATTCATGAATATTAGTGAATATCCCA 286
Db 245 ACATTTGCCACGAAAGCTGCGGATCCAGAAAGCAATTCATGAATGTTAGCGAAATCATCAA 304
QY 287 ACATCAAGGCTATCCCTGTGAGGAATATGAAGTCGCAACTGAAGTGGGTATATCCCTTTC 346
Db 305 ACACAGAGGTTATCCAGTGAGGATATGAAGTTGCAACGAGATGGGTATATCTTTC 364
QY 347 TGTTAACAGGATTTCTCGAGCCCTAGTGCAACCTTAAGAACAGAGTTTCCAGGCTGTGGT 406
Db 365 TGTGAACAGATTCCTCGGGGACAGACACGGTGTAAAGAAAGGAGGATCCAGGCCAGTGGT 424
QY 407 GTTACTGCGAGCATGCGCTAGTTGGAGGTGTAGCAACTGGATTTCCAACTGCCCAACAA 466
Db 425 GTTACTGCAACATGGTCTTCTTTGGGATGCTAGCAACTGGATTTCCAACTGCCCAACAA 484
QY 467 TAGCCTGGGCTTCATTCTTGGCAGATGCTGGTTTTCAGCTGTGGATGGGGAACACAGCGGG 526
Db 485 CAGCCTGGGTTTATTCTTGGCAGATGTCAGGTTTGTATGTGTGGATGGGGAACACAGCAGG 544
QY 527 AAACGCTGGTCTCGAAAACACAGACACTCTCCATAGACCAAGATGAGTTCTGGGCTTT 586
Db 545 AAACACCTGGTCTCGGAAGACCAAGACCCCTCTCCATAGATCAGATGAGTTCTGGGCTTT 604

587	QY	CAGTTATGATGAGATGGCTAGGTTTGACCTTCCTGCAGTGATAAACTTTATTTTGAGAA	646
605	Db	CAGTTATGATGAAATGGCTAGGTTTGACCTTCAGCTGTGATAAACTTTATCTTACAGAA	664
647	QY	ACGGGCGAGGAAAGATCTATTATGTGCGGTATTCCACAGGCGACCAACCATGGCTTTAT	706
665	Db	AACGGGCCAGAAAAGGTCTATTATGTGCGGTACTCACAGGCGACCAACCATGGCTTTAT	724
707	QY	TGCATTTTCCACCATCCAGAGCTGGCTCAGAAAATCAAAAATGATATTTTGTCTTTAGCAC	766
725	Db	TGCATTTTCCACCATCCAGAGCTAGCTCATAAAATCAAAAATGATATTTTGTCTTTAGCTC	784
767	QY	CATAGCCACTGTTAAGCATGCAAAAGCCCGGAGCAAAATTTTGTGTCGTCAGATAT	826
785	Db	TATAGCCACTGTTAAATATGCAAGAAAGTCTGATACCAAATTTCTGCTGTCGCAATAT	844
827	QY	GATGATCAAGGGATTTGTTTGGCAAAAAGAAATTTCTGTATCAGACAGATTTTCTCAGACA	886
845	Db	GATGATCAAGGTATTATTGGCAGACAGAGTTTATTACAGACTAGATTTTTCAGACA	904
887	QY	ACTTGTTATTACCTTTGTGGCCAGGTGATCTTGATTCAGATTTGTAGTAATATCATGTT	946
905	Db	GCTTTTTATTACCTTTGTGGCAGATGATCTTGACCAATCTGCAGCAACATCATCTT	964
947	QY	ACTTCTGGGTGGATTCAACACCAACAATATGAACATGAGCCGAGCAAGTGATATGCTGC	1006
965	Db	ACTCTGGGAGGTTTAAACAAACAATATGAACATGAGCAGAGCAATGTGTATGTCGC	1024
1007	QY	CCACATCTTGTGGAACAATCTGTGCAAAATATTTACTACTGGAGCCAGCAAGTGAATTC	1066
1025	Db	CCATACACCTGCTGGAAAGTCTGTGCAAGCAATTTCTCCACTGGAGCCAGCCGTGAATTC	1084
1067	QY	TGCTGAACCTCGGGCAATTTGACTGGGGAGTGAGACCAAAATCTGCAAAATGCAATCA	1126
1085	Db	TGGGGAACCTCGTGCTTTGACTGGGGAAGTGAGACCAAAATCAGAGAAATGTATATCA	1144
1127	QY	GCCAACTCTCTGAAGGTACAGAGTACAGAGATATGACGGTCCCTACAGCAATGTGGACAGG	1186
1145	Db	GCCAACTCTCTGAAGGTACAGAGTTCAGAGATATGATGTTCCCAACAGCAATGTGGACTGG	1204
1187	QY	AGGTCAGGACTGCTTCAAAATCCAGAGACGTGAAAATGCTGCTCTCTGAGGTGACCAA	1246
1205	Db	AGGTCAGGACTGCTTCAAAATCCAGATGATGTGAAAACATTAATTTCTGAAGTAACCAA	1264
1247	QY	CCTCATCTACCAAGAATAATTCCTGAATGGGCTCACGTGGATTTATCTGGGGTTTGGAA	1306
1265	Db	CCTCATCTACCAAGAACATTCCTGAGTGGGCTCATGTGGATTTATCTGGGGCTGGA	1324
1307	QY	TGCTCTCTACCGTATGTACAAATGAATATCATTCATCTGATCGAGAGAGAGACCAACCT	1366
1325	Db	TGCCCCCTCAGCGTGTTTACAAATGAAATATATACATCTGATGAAGC---AGGAGCCCAACCT	1381
1367	QY	TTCCCGAGGACGGTGTGAGGCGGTATTGTGAAGCATCTGCACATGACGATCTTTAGGACAA	1426
1382	Db	TCGCCAGGGAACCTGCGGGTCAAAATTGTGAGCTCTCTGCAA-----GTCTTAGAACAA	1435
1427	QY	CCTCTCTGAGGATGGG	1442
1436	Db	CCCATGCGGAATGGG	1451

RESULT 2

AK085719

LOCUS

DEFINITION

AK085719
Mus musculus 10 days lactation, adult female mammary gland cDNA,
RIKEN full-length enriched library, clone.D73026117
product:BA30415.1 (NOVEL LIPASE) (FRAGMENT) homolog [Homo sapiens],
full insert sequence.

ACCESSION

VERSION

KEYWORDS

SOURCE

AK085719
GI:26351760
HTC; CAP trapper.
Mus musculus (house mouse)

2678 bp

mRNA

linear

HTC 05-DEC-2002

Mus musculus
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus;
1
Carninci, P. and Hayashizaki, Y.
High-efficiency full-length cDNA cloning
Meth. Enzymol. 303, 19-44 (1999)
99279253
10349636
2
Carninci, P., Shibata, Y., Hayatsu, N., Sugahara, Y., Shibata, K.,
Itoh, M., Konno, H., Okazaki, Y., Muramatsu, M. and Hayashizaki, Y.
Normalization and subtraction of cap-trapper-selected cDNAs to
prepare full-length cDNA libraries for rapid discovery of new genes
Genome Res. 10 (10), 1617-1630 (2000)
20499374
11042159
3
Shibata, K., Itoh, M., Aizawa, K., Nagaoka, S., Sasaki, N., Carninci, P.,
Konno, H., Akiyama, J., Nishi, K., Kitsuai, T., Tashiro, H., Itoh, M.,
Sumi, N., Ishii, Y., Nakamura, S., Hazama, M., Nishida, T., Harada, A.,
Yamamoto, R., Matsumoto, H., Sakeguchi, S., Ikegami, T., Kashiwagi, K.,
Fujisawa, S., Inoue, K., Togawa, Y., Izawa, M., Chara, E., Watanabe, M.,
Yoneda, Y., Ishikawa, T., Ozawa, K., Tanaka, T., Matsura, S., Kawai, J.,
Okazaki, Y., Muramatsu, M., Inoue, Y., Kira, A. and Hayashizaki, Y.
RIKEN integrated sequence analysis (RISA) system - 384-format
sequencing pipeline with 384 multicapillary sequencer
Genome Res. 10 (11), 1757-1771 (2000)
20530913
11076861
4
Kawai, J., Shinagawa, A., Shibata, K., Yoshino, M., Itoh, M., Ishii, Y.,
Arakawa, T., Hara, A., Fukunishi, Y., Konno, H., Adachi, J., Fukuda, S.,
Aizawa, K., Izawa, M., Nishi, K., Kiyosawa, H., Kondo, S., Yamanaka, I.,
Saito, T., Okazaki, Y., Gjobori, T., Bono, H., Kasukawa, T., Saito, R.,
Kadota, K., Matsuda, H., Ashburner, M., Batalov, S., Casavant, T.,
Fleischmann, W., Gaasterland, T., Gless, C., King, B., Kochiwa, H.,
Kuehl, P., Lewis, S., Matsumoto, Y., Nikaido, I., Pesole, G.,
Quackenbush, J., Schiraldi, L. M., Staudt, F., Suzuki, R., Tomita, M.,
Wagner, L., Washio, T., Sakai, K., Okido, T., Furuno, M., Aono, H.,
Baldarelli, R., Batsch, G., Blake, J., Boffelli, D., Bojunga, N.,
Carninci, P., De Bonaldo, M. F., Brownstein, M. J., Bult, C.,
Fletcher, C., Fujita, M., Gariboldi, M., Gustincich, S., Hill, D.,
Hofmann, M., Hume, D. A., Kamiya, M., Lee, N. H., Lyons, P.,
Marchionni, L., Mashima, J., Mazzarelli, J., Momberti, P., Nordone, P.,
Ringo, B., Ringwald, M., Rodriguez, I., Sakamoto, N., Sasaki, H.,
Sato, K., Schonbach, C., Seya, T., Shibata, Y., Storch, K. F., Suzuki, H.,
Toyo-oka, K., Wang, K. H., Weitz, C., Whittaker, C., Wilming, L.,
Wyshak-Boris, A., Yoshida, K., Hasegawa, Y., Kawaji, H., Kohseki, S.,
Hayashizaki, Y.
Functional annotation of a full-length mouse cDNA collection
Nature 409 (6821), 695-690 (2001)
21085660
11217851
5
The PANTOM Consortium and the RIKEN Genome Exploration Research
Group Phase I & II Team.
Analysis of the mouse transcriptome based on functional annotation
of 60,770 full-length cDNAs
Nature 420, 563-573 (2002)
6 (bases 1 to 2678)
Adachi, J., Aizawa, K., Akimura, T., Arakawa, T., Bono, H., Carninci, P.,
Fukuda, S., Furuno, M., Hanagaki, T., Hara, A., Hashizume, W.,
Hayashida, K., Hayatsu, N., Hiramoto, K., Hiraoka, T., Hirozane, T.,
Hori, F., Imotani, K., Ishii, Y., Itoh, M., Kagawa, I., Kasukawa, T.,
Kato, H., Kawai, J., Kojima, Y., Kondo, S., Konno, H., Kouda, M.,
Koya, S., Kurihara, C., Matsuyama, T., Miyazaki, A., Murata, M.,
Nakamura, M., Nishi, K., Nomura, K., Numazaki, R., Oono, M., Ohsato, N.,
Okazaki, Y., Saito, R., Saitoh, H., Sakai, C., Sakai, K., Sakazume, N.,
Sano, H., Sasaki, D., Shibata, K., Shinagawa, A., Shiraki, T.,
Sugabe, Y., Tagami, M., Igawa, K., Takahashi, F., Takaku-Akahira, S.,
Takeda, Y., Tanaka, T., Tomaru, A., Toyata, T., Yasunishi, A.,
Muramatsu, M. and Hayashizaki, Y.

Direct Submission
Submitted (16-APR-2002) Yoshihide Hayashizaki, The Institute of Physical and Chemical Research (RIKEN), Laboratory for Genome Exploration Research Group, RIKEN Genomic Sciences Center (GSC), RIKEN Yokohama Institute, 1-7-22 Suehiro-cho, Teurumi-ku, Yokohama, Kanagawa 230-0045, Japan (E-mail: genome-res@gsr.riken.go.jp, URL: http://genome.gsc.riken.go.jp/, Tel: 81-45-503-9222, Fax: 81-45-503-9216)

cDNA library was prepared and sequenced in Mouse Genome Encyclopedia Project of Genome Exploration Research Group in Riken Genomic Sciences Center and Genome Science Laboratory in Riken. Division of Experimental Animal Research in Riken contributed to prepare mouse tissues.

Please visit our web site for further details.
URL: http://genome.gsc.riken.go.jp/
URL: http://fantom.gsc.riken.go.jp/
Location/Qualifiers
1..2678
/organism="Mus musculus"
/mol_type="mRNA"
/strain="C57BL/6J"
/db_xref="PANTOM.DB:D730026117"
/db_xref="taxon:10090"
/clone="D730026117"
/sex="female"
/tissue_type="mammary gland"
/clone_lib="RIKEN full-length enriched mouse cDNA library"
/dev_stages="10 days lactation, adult"
118..1386
/note="unnamed protein product; BA30415.1 (NOVEL LIPASE) (FRAGMENT) homolog [Homo sapiens] (SPTR|0951G2, evidence: PASTY, 88.5%ID, 100%length, match=1119)
putative"
/codon_start=1
/protein_id="BAC39517.1"
/db_xref="GI:26351761"
/translation="MSSEILSRVTVSHRVEILWILVAYLLQRNVNSGHLPTKADPE AFMVSEIILKIKVPSREVEVATEDYILSVNIPRGOTLRKKEGSRPVVLLQHLGLG DASWISNLNLSGLFADAGPDVWNGSGNTWSKHTLSLDDEFWAFSVDMA RFDPALNFILOKTKQKVVYVSGQTNGFIAPSTNPELAKIKYFALAPIATV KYRSPGKFLLLPDMIKVLFGRQEFYIYFRQFLYICGGMLDQICSNLILL GGFNTNNNSRANVYVAHTPAGTSVQNLHWSQVNSGELRAPDWSGTEKQKNO PTPIRYKVRDMVPTAMVTGGQDLNSPDDVKTLLSEVTNLIYHKNIPWAHVDYFING LDAPORVYNEIILHMKQEPNLPQGTCTCRVCL"
polya_signal
2657..2662
/note="putative"
2678
/note="putative"
polyA_site
BASE COUNT 795 a 522 c 555 g 806 t
ORIGIN

Query Match 49.2%; Score 1005.2; DB 11; Length 2678;
Best Local Similarity 84.0%; Pred No. 1.3e-152;
Matches 1173; Conservative 0; Mismatches 213; Indels 10; Gaps 3;

QY 48 TTTTAAACCCCAAAATTTCTTACTTTAGAAATAGTGTGTACATGGCAGGAA-AAAA 106
Db 39 TCTTTCAATCTCTGTTCTTCTTCTGACTCTAGATGATGTTTACATCTCTAGCAACAAAG 98

QY 107 TAAATGCAGATCTGGACCATCTGGAAACCTTGTCAAGACAGATGATGTCTCACAG 166
Db 99 CAAATGCCATGTGGACCATCTGAAATCTTGTCAAGAGTGTGATGTTTGCACAG 158

QY 167 AATGGAATGTGGCTTCTGATCTGGTGGCGATATGTTTCCAGAGAAATGTGAATTCAGT 226
Db 159 AGTGGAGATATGCTTCTGATCTGGTGGCGATATGTTTCTCCAAAGAAATGTGAATTCAGT 218

QY 227 ACATATGCCAATTAAGCTGTGGACCCAGAACATTCATGATATAGTGAATTCATCA 286
Db 219 ACATTTGCCACGAAGAGCTGGGATCCAGAGCATTCATGAATGTAGCGAAATCATCAA 278

QY 287 ACATCAAGGCTATCCCTGTGAGGAAATATGAAGTCGCAACTGAAGATGGGTATATCTTTTC 346

279 ACACAAGGGTTATCCAGTGGAGGATATGAAGTTGCAACGAGAGATGGGTATATCTTTC 338
QY 347 TGTAAACAGGATTTCTCGAGGCTAGTGCACACCTAAAGAGACAGTTCAGAGGCTGTGGT 406
Db 339 TGTAAACAGGATTTCTCGAGGCTAGTGCACACCTAAAGAGAGAGATTCAGGCGAGTGGT 398
QY 407 GTTACTGCAGCATGGCTTGTGGAGTGTAGCAACTGGATTTCCAACTGCTCCCAACAA 466
Db 399 GTTACTGCAGCATGGCTTGTGGAGTGTAGCAACTGGATTTCCAACTGCTCCCAACAA 458
QY 467 TGTACTGGCTTCTTCTCGAGGCTAGTGCACACCTAAAGAGAGAGATTCAGGCGAGTGGT 526
Db 459 TGTACTGGCTTCTTCTCGAGGCTAGTGCACACCTAAAGAGAGAGATTCAGGCGAGTGGT 518
QY 527 TGTACTGGCTTCTTCTCGAGGCTAGTGCACACCTAAAGAGAGAGATTCAGGCGAGTGGT 586
Db 519 TGTACTGGCTTCTTCTCGAGGCTAGTGCACACCTAAAGAGAGAGATTCAGGCGAGTGGT 578
QY 587 TGTACTGGCTTCTTCTCGAGGCTAGTGCACACCTAAAGAGAGAGATTCAGGCGAGTGGT 646
Db 579 TGTACTGGCTTCTTCTCGAGGCTAGTGCACACCTAAAGAGAGAGATTCAGGCGAGTGGT 638
QY 647 TGTACTGGCTTCTTCTCGAGGCTAGTGCACACCTAAAGAGAGAGATTCAGGCGAGTGGT 706
Db 639 TGTACTGGCTTCTTCTCGAGGCTAGTGCACACCTAAAGAGAGAGATTCAGGCGAGTGGT 698
QY 707 TGTACTGGCTTCTTCTCGAGGCTAGTGCACACCTAAAGAGAGAGATTCAGGCGAGTGGT 766
Db 699 TGTACTGGCTTCTTCTCGAGGCTAGTGCACACCTAAAGAGAGAGATTCAGGCGAGTGGT 758
QY 767 TGTACTGGCTTCTTCTCGAGGCTAGTGCACACCTAAAGAGAGAGATTCAGGCGAGTGGT 826
Db 759 TGTACTGGCTTCTTCTCGAGGCTAGTGCACACCTAAAGAGAGATTCAGGCGAGTGGT 818
QY 827 TGTACTGGCTTCTTCTCGAGGCTAGTGCACACCTAAAGAGAGATTCAGGCGAGTGGT 886
Db 819 TGTACTGGCTTCTTCTCGAGGCTAGTGCACACCTAAAGAGAGATTCAGGCGAGTGGT 878
QY 887 TGTACTGGCTTCTTCTCGAGGCTAGTGCACACCTAAAGAGAGATTCAGGCGAGTGGT 946
Db 879 TGTACTGGCTTCTTCTCGAGGCTAGTGCACACCTAAAGAGAGATTCAGGCGAGTGGT 938
QY 947 TGTACTGGCTTCTTCTCGAGGCTAGTGCACACCTAAAGAGAGATTCAGGCGAGTGGT 1006
Db 939 TGTACTGGCTTCTTCTCGAGGCTAGTGCACACCTAAAGAGAGATTCAGGCGAGTGGT 998
QY 1007 TGTACTGGCTTCTTCTCGAGGCTAGTGCACACCTAAAGAGAGATTCAGGCGAGTGGT 1066
Db 999 TGTACTGGCTTCTTCTCGAGGCTAGTGCACACCTAAAGAGAGATTCAGGCGAGTGGT 1058
QY 1067 TGTACTGGCTTCTTCTCGAGGCTAGTGCACACCTAAAGAGAGATTCAGGCGAGTGGT 1126
Db 1059 TGTACTGGCTTCTTCTCGAGGCTAGTGCACACCTAAAGAGAGATTCAGGCGAGTGGT 1118
QY 1127 TGTACTGGCTTCTTCTCGAGGCTAGTGCACACCTAAAGAGAGATTCAGGCGAGTGGT 1186
Db 1119 TGTACTGGCTTCTTCTCGAGGCTAGTGCACACCTAAAGAGAGATTCAGGCGAGTGGT 1178
QY 1187 TGTACTGGCTTCTTCTCGAGGCTAGTGCACACCTAAAGAGAGATTCAGGCGAGTGGT 1246
Db 1179 TGTACTGGCTTCTTCTCGAGGCTAGTGCACACCTAAAGAGAGATTCAGGCGAGTGGT 1238
QY 1247 TGTACTGGCTTCTTCTCGAGGCTAGTGCACACCTAAAGAGAGATTCAGGCGAGTGGT 1306
Db 1239 TGTACTGGCTTCTTCTCGAGGCTAGTGCACACCTAAAGAGAGATTCAGGCGAGTGGT 1298
QY 1307 TGTACTGGCTTCTTCTCGAGGCTAGTGCACACCTAAAGAGAGATTCAGGCGAGTGGT 1366
Db 1299 TGTACTGGCTTCTTCTCGAGGCTAGTGCACACCTAAAGAGAGATTCAGGCGAGTGGT 1355
QY 1367 TGTACTGGCTTCTTCTCGAGGCTAGTGCACACCTAAAGAGAGATTCAGGCGAGTGGT 1426
Db 1356 TGTACTGGCTTCTTCTCGAGGCTAGTGCACACCTAAAGAGAGATTCAGGCGAGTGGT 1409

167 AATGGAAATGCTGCTCTGATCTGCTGGGTATATGTTCCAGAGAAATGTGAATTCAGT 226
 Db |||||
 167 AGTGGAGATATGGCTCTGATCTGCTGGGTATATGTTCCAGAGAAATGTGAATTCAGT 226
 Qy |||||
 227 ACATATGCCAACTAAAGCTGTGACCCAGAGCAATCATGAATATTAGTGAATCATCA 286
 Db |||||
 227 ACATTTGCCACGAAGCTGGATCCAGAGCAATCATGAATATTAGTGAATCATCA 286
 Qy |||||
 287 ACATCAAGGCTATCCCTGTCAGGAATATGAATGCGCAACTGAAGATGGGTATATCTTTC 346
 Db |||||
 287 ACAGAGGTTATCCCAAGTGAAGATGAAGTTGCAACGAAGATGGGTATCTTCTTTC 346
 Qy |||||
 347 TGTTAACAGGATTCCTCGAGCCCTPAGTGCACCTTAAGACAGAGTTCCAGGCTGTGT 406
 Db |||||
 347 TGTGAACAGAAATCCCTCGGAGCAGACACAGTGTAAAGAGGAAGATCCAGGCGAGTGT 406
 Qy |||||
 407 GTTACTCGCATGCTGCTAGTTGAGGTGTAGCACTGCACTGGATTTCCAACTGCCCAACAA 466
 Db |||||
 407 GTTACTCGCAATGCTGCTTCTGGGATGCTAGCACTGGATTTCCAACTGCCCAACAA 466
 Qy |||||
 467 TAGCCTGGGCTTCATCTCGCAGATGCTGTTTGTGACGTGTGATGGGGAACAGCAGGG 526
 Db |||||
 467 CAGCCTGGGTTTATCTCGCAGATGCTGTTTGTGATGCTGATGGGGAACAGCAGAGG 526
 Qy |||||
 527 AAACGCTGCTGCTGAAAACACAGACACTCTCCATAGACAGATGAGTCTGGGCTTT 586
 Db |||||
 527 AAACACCTGCTGCTGGAAGCAGACAGCCCTCTCCATAGATCAAGATGAGTCTGGGCTTT 586
 Qy |||||
 587 CAGTTATGATGAGATGCTAGTGTGCTGCTTCTGCTGAGTGAATTAATTTTTCAGAA 646
 Db |||||
 587 CAGTTATGATGATGCTAGTGTGCTGCTTCTGCTGAGTGAATTAATTTTTCAGAA 646
 Qy |||||
 647 AACGGCCAGAAAAGATCTATTATGCTGCTATTCACAGGCAACCACTGGCTTTAT 706
 Db |||||
 647 AACGGCCAGAAAAGTCTATTATGCTGCTATTCACAGGCAACCACTGGCTTTAT 706
 Qy |||||
 707 TGCAATTTTCCACCTGCGCAGAGCTGGCTCAGAAAATCAAAATGATTTTGTCTTAGCAC 766
 Db |||||
 707 TGCAATTTTCCACCTGCGCAGAGCTAGCTCATAAAATCAAAATGATTTTGTCTTAGCTC 766
 Qy |||||
 767 CATAGCCTGTTAAGCATGCAAAAGCCCGGAGCAAAATTTTGTGCTGCCAGATAT 826
 Db |||||
 767 CATAGCCTGTTAAGCATGCAAAAGTCTGCTGCAAAATTTTGTGCTGCCAGATAT 826
 Qy |||||
 827 GATGATCAAGGATTTGTCGCAAAAGAAATTTCTGATCAGACCAAGATTTCTCAGACA 886
 Db |||||
 827 GATGATCAAGGATTTATTTGCGCAGACAGAGTTTATACCAAGACTAGATTTTTCAGACA 886
 Qy |||||
 887 ACTGTTATTTACCTTTGTCGCGAGGATTTCTGATCAGATTTGTAGTAATCATGTT 946
 Db |||||
 887 GCTTTTATTTACCTTTGTCGCGAGGATTTCTGATCAGATTTGTAGTAATCATGTT 946
 Qy |||||
 947 ACTTCTGGGTGATTTCAACACCAACAAATATGAATGAGCGGAGCAAGTGTATATGCTGC 1006
 Db |||||
 947 ACTTCTGGGTGATTTTACACCAACAAATATGAATGAGCGGAGCAAGTGTATATGCTGC 1006
 Qy |||||
 1007 CCACACTCTGTGTGAAACATCTGTGCAAAATTTCTGATCAGTGGAGCGAGGAGTGAATTC 1066
 Db |||||
 1007 CCATACACTCTGTGTGAAACATCTGTGCAAAATTTCTGATCAGTGGAGCGGAGTGAATTC 1066
 Qy |||||
 1067 TGTGTAACCTCGGGGATTTGATCGGGGAGTGAGACCAAAATCTGGAATAATGCAATCA 1126
 Db |||||
 1067 TGTGTAACCTCGGGGATTTGATCGGGGAGTGAGACCAAAATCTGGAATAATGCAATCA 1126
 Qy |||||
 1127 GCCAACTCTGTGATGATCAGATGATGATGATGATGATGATGATGATGATGATGATGATG 1186
 Db |||||
 1127 GCCAACTCTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 1186
 Qy |||||
 1187 AGGTGAGGACTGCTCTTCAATCCAGAGACGTGAATGCTGCTCTGAGGTGACCAA 1246
 Db |||||
 1187 AGGTGAGGACTGCTCTTCAATCCAGAGATGATGATGATGATGATGATGATGATGATGATG 1246

1247 CCTCATCTACCATGAATATTTCTGAATGGGTCTCAGTGGATTTTCATCTGGGGTTTGA 1306
 Db |||||
 1247 CCTCATCTACCATGAATATTTCTGAATGGGTCTCAGTGGATTTTCATCTGGGGTTTGA 1306
 Qy |||||
 1307 TGCTCTCACCGTATGATCAATGAATCATCACTCTGATCGACGAGGAGGAGACCACT 1366
 Db |||||
 1307 TGCCCTCTCAGGCTTTTACCAATGAATTTATACATCTGATGAAGC---AGAGGCCAACCT 1363
 Qy |||||
 1367 TTCCAGGAGCGGTGTGAGGCCGTATTGTGAAGCATCTGACACTGACGATCTTAGGACAA 1426
 Db |||||
 1364 TCCCGAGGACCTGCGAGGTCAATTTGTGAAGCTCTGCAA-----GTCTTAGAACAA 1417
 Qy |||||
 1427 CTTCTGAGGAGTGGG 1442
 Db |||||
 1418 CCCCATGCGGAATGGG 1433

RESULT 4

AK019504

LOCUS DEFINITION

Mus musculus 0 day neonate skin cDNA, RIKEN full-length enriched library, clone:4632427C23 product:BA30415.1 (NOVEL LIPASE) (FRAGMENT) homolog [Homo sapiens], full insert sequence.

ACCESSION 2

VERSION

KEYWORDS

SOURCE

ORGANISM

REFERENCE

AUTHORS

TITLE

JOURNAL

MEDLINE

PUBMED

REFERENCE

AUTHORS

TITLE

JOURNAL

MEDLINE

PUBMED

REFERENCE

AUTHORS

TITLE

JOURNAL

MEDLINE

PUBMED

REFERENCE

AUTHORS

TITLE

JOURNAL

MEDLINE

PUBMED

REFERENCE

AUTHORS

TITLE

JOURNAL

MEDLINE

PUBMED

REFERENCE

AUTHORS

TITLE

JOURNAL

MEDLINE

PUBMED

REFERENCE

AUTHORS

TITLE

JOURNAL

MEDLINE

PUBMED

REFERENCE

AUTHORS

TITLE

JOURNAL

MEDLINE

PUBMED

REFERENCE

2927 bp mRNA linear HTC 05-DEC-2002
 Mus musculus 0 day neonate skin cDNA, RIKEN full-length enriched library, clone:4632427C23 product:BA30415.1 (NOVEL LIPASE) (FRAGMENT) homolog [Homo sapiens], full insert sequence.

AK019504.1 GI:12859754

HTC; CAP trapper.

Mus musculus (house mouse)

Mus musculus

Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.

Carninci, P. and Hayashizaki, Y.

High-efficiency full-length cDNA cloning

Meth. Enzymol. 303, 19-44 (1999)

99279253

10349636

Carninci, P., Shibata, Y., Hayatsu, N., Sugahara, Y., Shibata, K., Itoh, M., Konno, H., Okazaki, Y., Muramatsu, M. and Hayashizaki, Y.

Normalization and subtraction of cap-trapper-selected cDNAs to prepare full-length cDNA libraries for rapid discovery of new genes

Genome Res. 10 (10), 1617-1630 (2000)

20493374

11042159

Shibata, K., Itoh, M., Aizawa, K., Nagaoka, S., Sasaki, N., Carninci, P., Konno, H., Akiyama, J., Nishi, K., Katsunai, T., Tashiro, H., Itoh, M., Sumi, N., Ishii, Y., Nakamura, S., Hazama, M., Nishino, T., Harada, A., Yamamoto, R., Matsumoto, H., Sakaguchi, S., Ikegami, T., Kashiwagi, K., Fujiwara, S., Inoue, K., Togawa, Y., Izawa, M., Ohara, E., Watahiki, M., Yoneda, Y., Ishikawa, T., Ozawa, K., Tanaka, T., Matsura, S., Kawai, J., Okazaki, Y., Muramatsu, M., Inoue, Y., Kira, A. and Hayashizaki, Y.

RIKEN integrated sequence analysis (RISA) system--384-format sequencing pipeline with 384 multicapillary sequencer

Genome Res. 10 (11), 1757-1771 (2000)

20530913

11076861

Kawai, J., Shinagawa, A., Shibata, K., Yoshino, M., Itoh, M., Ishii, Y., Arakawa, T., Hara, A., Fukunishi, Y., Konno, H., Adachi, J., Fukuda, S., Aizawa, K., Izawa, M., Nishi, K., Kiyosawa, H., Kondo, S., Yamanaka, I., Saito, T., Okazaki, Y., Gojobori, T., Bono, H., Kasukawa, T., Saito, R., Kadota, K., Matsuda, H., Ashburner, M., Batalov, S., Casavant, T., Fleischmann, W., Gaasterland, T., Gissi, C., King, B., Kochiwa, H., Kuehl, P., Lewis, S., Matsuo, Y., Nikaide, I., Pesole, G., Quackenbush, J., Schriml, L. M., Staubli, F., Suzuki, R., Tomita, M., Wagner, L., Washio, T., Sakai, K., Okido, T., Furuno, M., Aono, H., Baldarelli, R., Barsh, G., Blake, J., Boffelli, D., Bojunga, N., Carninci, P., de Bonaldo, M. F., Brownstein, M. J., Buit, C., Fletcher, C., Fujita, M., Gariboldi, M., Gustincich, S., Hill, D., Hofmann, M., Hume, D. A., Kamiya, M., Lee, N. H., Lyons, P., Marchionni, L., Mashima, J., Mazzarelli, J., Mombaerts, P., Nordone, P., Ring, B., Ringwald, M., Rodriguez, I., Sakamoto, N., Sasaki, H.,


```

SOURCE      Homo sapiens (human)
ORGANISM    Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE   1 (bases 1 to 629)
            NIH-MGC http://mgi.nci.nih.gov/.
            National Institutes of Health, Mammalian Gene Collection (MGC)
            Unpublished
COMMENT     Contact: Robert Strausberg, Ph.D.
            Email: csapbs-remail.nih.gov
            Tissue Procurement: James Cleaver, M.D.
            cDNA Library Preparation: Life Technologies, Inc.
            cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL) DNA
            Sequencing by: Incyte Genomics, Inc.
            Clone distribution: MGC clone distribution information can be
            found through the I.M.A.G.E. Consortium/LLNL at:
            http://image.llnl.gov
            Plate: L1AM10631 row: i column: 12
            High quality sequence stop: 629.
FEATURES    source
            1..629
            /organism="Homo sapiens"
            /mol_type="mRNA"
            /db_xref="taxon:9606"
            /clone="IMAGE:4777427"
            /lab_host="DH10B (T1 phage-resistant)"
            /clone_lib="NCI CGAP Skn3"
            /note="Organ: skin; Vector: pCMV-SPORT6; Site: 1: NotI;
            Site 2: SalI; Cloned unidirectionally. Primer: Oligo dt.
            Average insert size 1.5kb. Library constructed by Life
            Technologies. Note: this is a NCI CGAP Library."
BASE COUNT  204 a 118 c 107 g 200 t
ORIGIN
Query Match      29.0%; Score 592; DB 10; Length 629;
Best Local Similarity 100.0%; Pred. No. 9.3e-86;
Matches 592; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 1426 ACCTCCTGAGGAGTGGGCTAGGACCATGAGCAGAAATACGGAGAGAGACCTAG 1485
DB 629 ACCTCCTGAGGAGTGGGCTAGGACCATGAGCAGAAATACGGAGAGAGACCTAG 570
QY 1486 TATACATTTTTCAGATTCCTGTCGACTGGCAGTAAATCCGACACTTACATTTT 1545
DB 569 TATACATTTTTCAGATTCCTGTCGACTGGCAGTAAATCCGACACTTACATTTT 510
QY 1546 TTTCTGTAATTAAGTACTTTATAGTAATAGAGTTTGTATGCTATATATTTCT 1605
DB 509 TTTCTGTAATTAAGTACTTTATAGTAATAGAGTTTGTATGCTATATATTTCT 450
QY 1606 ACCATCTTGAAGGGTAGGTTTACCTGATAGCCAGAAAATATCTAGACATTTCTATATC 1665
DB 449 ACCATCTTGAAGGGTAGGTTTACCTGATAGCCAGAAAATATCTAGACATTTCTATATC 390
QY 1666 ATTGAGTAAATCTTTTAAACACCTATTTGTTTCTATAGCCATATTTTGGAGCA 1725
DB 389 ATTGAGTAAATCTTTTAAACACCTATTTGTTTCTATAGCCATATTTTGGAGCA 330
QY 1726 CTAAAGTAAATGGCAATTCGGACAGATATTGAGGCTGAGCTGTGATATTGTTG 1785
DB 329 CTAAAGTAAATGGCAATTCGGACAGATATTGAGGCTGAGCTGTGATATTGTTG 270
QY 1786 ACTTTGACAAAATAGCTAGACATTTTCACTTGTGTCACAGAGACATTAACACTACCTC 1845
DB 269 ACTTTGACAAAATAGCTAGACATTTTCACTTGTGTCACAGAGACATTAACACTACCTC 210
QY 1846 AGGAAGCTGAGCTGCTTTAAGGACACCAACCAACAAATCATGTTTACGATGATGATAA 1905
DB 209 AGGAAGCTGAGCTGCTTTAAGGACACCAACCAACAAATCATGTTTACGATGATGATAA 150
QY 1906 TCTATGTTAAGCATTTCTCAGAAATAGGCCAAAGTTTATAGTTGATCTCAGGGAGAAA 1965
DB 149 TCTATGTTAAGCATTTCTCAGAAATAGGCCAAAGTTTATAGTTGATCTCAGGGAGAAA 90

```

```

QY 1966 TTTTATAGATCTTTATGAGTTTCCATTAATGATCTGCTGATTTACATATA 2017
DB 89 TTTTATAGATCTTTATGAGTTTCCATTAATGATCTGCTGATTTACATATA 38

```

RESULT 6
BX103621/c
LOCUS BX103621 Soares fetal lung NbHL19W Homo sapiens cDNA clone
DEFINITION IMAGE:307244, mRNA sequence.
ACCESSION BX103621
VERSION BX103621.1 GI:27845805
KEYWORDS EST.
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
1 (bases 1 to 556)
Ebert, U., Heil, O., Hennig, S., Neubert, P., Partsch, E., Peters, M.,
Radelof, U., Schneider, D. and Korn, B.
Human UnigeneSet - RZPD3
Unpublished
Contact: Ina Rolfs
RZPD Deutsches Ressourcenzentrum fuer Genomforschung GmbH
Im Neuenheimer Feld 580, D-69120 Heidelberg, Germany
RZPD; IMAGE:998G21690.
RZPDLIB; I.M.A.G.E. cDNA Clone Collection;
Human UnigeneSet - RZPD3 (RZPDLIB No.972)
bin/showLib.pl.cgi?response?libNo=972 Contact: Ina Rolfs
RZPD Deutsches Ressourcenzentrum fuer Genomforschung GmbH
Heubnerweg 6, D-14059 Berlin, Germany
Tel: +49 30 32639 101
Fax: +49 30 32639 111
www.rzpd.de
This clone is available royalty-free from RZPD;
contact RZPD (clone@rzpd.de) for further information. Seq primer:
M13r, Primer sequence: TTTACACAGGAACAGCATGAC.
FEATURES
source
1..556
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/clone="IMAGE:998G21690 ; IMAGE:307244"
/dev_stage="19 weeks"
/lab_host="DH10B (ampicillin resistant)"
/clone_lib="Soares fetal lung NbHL19W"
/note="Organ: lung; Vector: pT73D (Pharmacia) with a
modified polylinker; Site: 1: Not I; Site 2: Eco RI; 1st
strand cDNA was primed with a Not I - oligo(dT) primer
[5'-TGTTACCAATCTGAAGTGGGAGCGGCGCAATTTTTTTTTTTT-3']
double-stranded cDNA was size selected, ligated to Eco RI
adapters (Pharmacia), digested with Not I and cloned into
the Not I and Eco RI sites of a modified pT73 vector
(Pharmacia). Library went through one round of
normalization to a Cot = 5. Library constructed by Bento
Soares and M.Patima Bonaldi. This library was constructed
from the same fetus as the fetal heart library. Soares
fetal heart NbHL19W."
BASE COUNT 177 a 106 c 93 g 179 t 1 others
ORIGIN
Query Match 26.6%; Score 543; DB 13; Length 556;
Best Local Similarity 99.6%; Pred. No. 7.8e-78;
Matches 554; Conservative 0; Mismatches 1; Indels 1; Gaps 1;
QY 1425 AACCTCCTGAGGATGGGGCTAGGACCATGAGCAGAAATACGGAGAGAGACCTA 1484
DB 556 AACCTCCTGAGGATGGGGCTAGGACCATGAGCAGAAATACGGAGAGAGACCTA 497
QY 1485 GTATACATTTTTCAGATTCCTCTGCACTTGGCACTAAATCCGACACTTACATTTT 1544

496 GTATACATTTTTCAGATTCCTGCACTTGGCAGCACTAAATCCGACACTTACATTTTACATTTT 437
 1545 TTTTCTGTAATTTAAAGTACTTATTAGGTAATAGAGGTTTGTGATGCTATTATATATTC 1604
 436 TTTTCTGTAATTTAAAGTACTTATTAGGTAATAGAGGTTTGTGATGCTATTATATATTC 377
 1605 TACCATCTTGAAGGGTGGTGTATCTGATAGCCAGAAATATCTAGACATTTCTCTATAT 1664
 376 TACCATCTTGAAGGGTGGTGTATCTGATAGCCAGAAATATCTAGACATTTCTCTATAT 317
 1665 CATTCAGGTAATCTCTTTAAACACCTATTCTTTTCTAATAAGCCATATTTTGGAGC 1724
 316 CATTCAGGTAATCTCTTTAAACACCTATTCTTTTCTAATAAGCCATATTTTGGAGC 257
 1725 ACTAAAGTAAATGGCAATTTGGACAGATATTGGAGTCTGGAGTCTGTGGATTATGTT 1784
 256 ACTAAAGTAAATGGCAATTTGGACAGATATTGGAGTCTGGAGTCTGTGGATTATGTT 197
 1785 GACTTTGACAAATAGCTAGACATTTTCACTCTGTTGCCAGAGACATTAACATCTT 1844
 196 GACTTTGACAAATAGCTAGACATTTTCACTCTGTTGCCAGAGACATTAACATCTT 137
 1845 CAGGAAGCTGAGCTGCTTTAAGGACAAACCAACCAAAATCACTGTTTACAGTATGATGAA 1904
 136 CAGGAAGCTGAGCTGCTTTAAGGACAAACCAACCAAAATCACTGTTTACAGTATGATGAA 77
 1905 ATCTATGTAAGCAATCTCAGATAAGGCCAAGTTTATAGTTGAT-CTAGAGGAAGAA 1963
 76 ATCTATGTAAGCAATCTCAGATAAGGCCAAGTTTATAGTTGATGATCTAGAGGAAGAA 17
 1964 AATTTTATAGGATGTT 1979
 16 AATTTTATAGGATGTT 1

RESULT 7
 BF135102
 LOCUS 601780142F1 NCI_CGAP_Lu30 Mus musculus cDNA clone IMAGE:4008131 5',
 DEFINITION mRNA sequence.
 ACCESSION BF135102
 VERSION BF135102.1 GI:10974142
 KEYWORDS EST.
 SOURCE Mus musculus (house mouse)
 ORGANISM Mus musculus
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
 NIH-MGC http://mgi.nci.nih.gov/
 National Institutes of Health, Mammalian Gene Collection (MGC)
 Unpublished
 Contact: Robert Strausberg, Ph.D.
 Email: cgabs-r@mail.nih.gov
 Tissue Procurement: Gilbert Smith, Ph.D.
 CDNA Library Preparation: Life Technologies, Inc.
 DNA Sequencing By: Incyte Genomics, Inc.
 Clone Distribution: NCI-CGAP clone distribution information can be
 found through the I.M.A.G.E. Consortium/LLNL at:
 http://image.llnl.gov
 Plate: LAM9242 row: c column: 12
 High quality sequence stop: 695.
 Location/Qualifiers
 1. .954
 /organism="Mus musculus"
 /mol_type="mRNA"
 /strain="CZECH II"
 /db_xref="taxon:10090"
 /clone="IMAGE:4008131"
 /tissue_type="tumor, metastatic to mammary"
 /lab_host="PH10B"
 /clone_lib="NCI CGAP Lu30"
 /note="Organ: lung; Vector: pCMV-SPORT6; Site_1: NotI;

BASE COUNT 280 a 222 c 222 g 230 t
 ORIGIN

Query Match 24.7%; Score 505; DB 10; Length 954;
 Best Local Similarity 85.3%; Pred. No. 8.4e-72;
 Matches 598; Conservative 0; Mismatches 100; Indels 3; Gaps 3;

Qy 618 CTGCGAGTGAATACTTTATTTTGCAGAAACCGGCGAGGAAAGATCTATTATGTCGGC 677
 Db 1 CCAGCTGTGATAAATCTTTATTTACAGAAACCGGCGAGGAAAGATCTATTATGTCGGC 60
 Qy 678 TATTACAGGCGCACCATCGGCTTTATTGCAATTTTCCACCATGCGAGAGCTGGCTCAG 737
 Db 61 TACTCAGAGGCGCACCATCGGCTTTATTGCAATTTTCCACCATGCGAGAGCTAGTCTAT 120
 Qy 738 AAAATCAAAATGATTTTGTGCTAGCAATGATGATGATGATGATGATGATGATGATGAT 797
 Db 121 AAAATCAAAATGATTTTGTGCTAGCTCTTATGCAATGATGATGATGATGATGATGAT 180
 Qy 798 GGGACCAAAATTTTGTGCTGCGAGATATGATGATGATGATGATGATGATGATGATGAT 857
 Db 181 GGTACCAAAATTTGCTGCTGCGAGATATGATGATGATGATGATGATGATGATGATGAT 240
 Qy 858 TTTCTGTATCAGACCAAGATTTTCTCAGCAATTTGTTTATTTATTTTGTGCTGCGAGAT 917
 Db 241 TTTTATACCAAGATTTTCTCAGCAATTTTGTGCTGCGAGATTTTATTTATTTTATTTAC 299
 Qy 918 CTTGATCAGATTTGTATATATGATGATGATGATGATGATGATGATGATGATGATGATGAT 977
 Db 300 CTGACCAAAATCTGCGAGCAATCTTCTTCTGCGAGATTTTATTTATTTTATTTATTTAT 359
 Qy 978 AACATGAGCGGAGCAAGTGTATATGCTGCCCACTCTTCTGCGAGATTTTATTTATTTATTT 1037
 Db 360 AACATGAGCGGAGCAAGTGTATATGCTGCCCACTCTTCTGCGAGATTTTATTTATTTATTT 419
 Qy 1038 ATTCTACACTGAGCGGAGCAAGTGTATGCTGCGAGATTTTATTTATTTATTTATTTATTT 1097
 Db 420 ATTCTACACTGAGCGGAGCAAGTGTATGCTGCGAGATTTTATTTATTTATTTATTTATTT 479
 Qy 1098 GAGACCAAAATCTGCGAGCAATCTTCTGCGAGATTTTATTTATTTATTTATTTATTTATTT 1157
 Db 480 GAGACCAAAATCTGCGAGCAATCTTCTGCGAGATTTTATTTATTTATTTATTTATTTATTT 539
 Qy 1158 ATGACGCTCCCTCAGCAATGTCGAGAGGTCAGAGTCTGCGAGATTTTATTTATTTATTTATTT 1217
 Db 540 ATGATGCTCCCTCAGCAATGTCGAGAGGTCAGAGTCTGCGAGATTTTATTTATTTATTTATTT 598
 Qy 1218 GTGAAATGCTGCTCTCTGAGGTGACCACTCTATCTTACCATAGATATTTCTGATGATG 1277
 Db 599 GTGAAATGCTGCTCTCTGAGGTGACCACTCTATCTTACCATAGATATTTCTGATGATG 658
 Qy 1278 GCTCACTGATTTCTATCTGCGGTTTGGATGCTCTCAGCG 1318
 Db 659 GGTCACTGGA-TTCATCTGCGGTTTGGATGCTCTCAGCG 698

RESULT 8
 BF141567
 LOCUS 601788654F1 NCI_CGAP_Lu30 Mus musculus cDNA clone IMAGE:4016413 5',
 DEFINITION mRNA sequence.
 ACCESSION BF141567
 VERSION BF141567.1 GI:10980607
 KEYWORDS EST.
 SOURCE Mus musculus (house mouse)
 ORGANISM Mus musculus
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
 REFERENCE 1 (bases 1 to 1079)

AUTHORS NIH-MGC <http://mgc.nci.nih.gov/>.
TITLE National Institutes of Health, Mammalian Gene Collection (MGC)
JOURNAL Unpublished
COMMENT Contact: Robert Strausberg, Ph.D.
 Email: cgapbs-remail.nih.gov
 Tissue Procurement: Gilbert Smith, Ph.D.
 cDNA Library Preparation: Life Technologies, Inc.
 cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)
 DNA Sequencing by: Incyte Genomics, Inc.
 Clone distribution: NCI-CGAP clone distribution information can be found through the I.M.A.G.E. Consortium/LLNL at:
<http://image.llnl.gov>
 Plate: LHAM9263 row: 1 column: 14
 High quality sequence stop: 663.
 Location/Qualifiers
 1. 1079

FEATURES

source
 /organism="Mus musculus"
 /mol_type="mRNA"
 /strain="CZECH II"
 /db_xref="taxon:10090"
 /clone="IMAGE:4016413"
 /tissue_type="tumor, metastatic to mammary"
 /lab_host="DH10B"
 /clone_lib="NCI CGAP Lu30"
 /note="Organ: lung; Vector: pCMV-SPORT6; Site 1: NotI; Site 2: SalI; transgenic model WNT-1, expression driven by MMV-LTR enhancer; Cloned unidirectionally. Primer: Oligo dt. Library constructed by Life Technologies. Investigator providing samples: Gilbert Smith, NIH"
 BASE COUNT 312 a 267 C 263 G 237 T
 ORIGIN
 Query Match 24.3%; Score 497.6; DB 10; Length 1079;
 Best Local Similarity 84.1%; Pred. No. 1.2e-70;
 Matches 621; Conservative 0; Mismatches 109; Indels 8; Gaps 5;
 QY 381 AAGAGACAGGTTCCAGGCTTGTGTGTTACTGACGATGCGCTAGTTGGAGTGTCTAGC 440
 DB 20 AAGAGGAGGATCCAGGCCAGTGTGTACTGCAACATGCTCTTCTGGGATGCTAGC 79
 QY 441 AACTGGATTCCACCTGCCCAACATAGCTGGCTTCACTTGGCAGATGCTGTTT 500
 DB 80 AACTGGATTCCACCTGCCCAACATAGCTGGCTTCACTTGGCAGATGCTGTTT 139
 QY 501 GACGTGTGGATGGGAAACAGCAGCGGAAACGCC-TGGTCTCGMAAAACACAGACACTCTC 559
 DB 140 GATGTGTGGATGGGAAACAGCAGAGGAAACACCGTGTCTCGGAGCACAGACCTCTC 199
 QY 560 CATGACCAAGATGAGTCTGGGCTTTCAGTTATGATGAGATGCTAGTTGACCTTCC 619
 DB 200 CATGATCAAGATGAGTCTGGGCTTTCAGTTATGATGAGATGCTAGTTGACCTTCC 259
 QY 620 TGCAGTGATAAATCTTATTTTGCAGAAAACGGGCCAGGAAAGATCTATTATGCGGCTA 679
 DB 260 AGCTGTGATAAATCTTATTTACAGAAAACGGGCCAGGAAAGATCTATTATGCGGCTA 319
 QY 680 TTCACAGGCAACACCATGGGCTTTATGCAATTTTCCACCATGCGCAGCTGGCTCAGAA 739
 DB 320 CTCACAGGGCCACCATGGGCTTTATTTGCAATTTTCCACCATGCGCAGCTAGCTCATAA 379
 QY 740 AATCAAAATGATTATTGTTGTTAGCACCCATAGCCACTGTTAAGCATCAAAAAGCCCGG 799
 DB 380 AATCAAAATGATTATTGTTGTTAGCTCTTATGACCACTGTTAATATGCAAGATCTCTG 439
 QY 800 GACCAAAATTTTGTGTCGAGATATGATCAAGGATGTTTGGCAAAAAGAAATT 859
 DB 440 TACCAAAATTTCTGCTGCTGCCAGATATGATGATCAAGGATTTATTTGGCAGACAAGATT 499
 QY 860 TCTGTATCAGACCAATTTCTCAGACACTGTTGTTATTTACCTTTGTGCCAGGATTTCT 919
 DB 500 TTTATACCAAGTATTTTTCAGGACGCTTTTATTTACCTTTGTGCCAGGATTTCT 559
 QY 920 TGATCAGATTTGTAGTAATATCATGTTATCTTTCTGGGTGGATTTCAACACCAACATATGAA 979

Db 560 CGACAAATCTGCAGCAACATCATCTTACTCTGGGAGGTTTAAACACAAACCATATGAA 619
 QY 980 CATGAGCCGAGCA--AGTGTATATGCTGCCACACTCTTGTCTGGAACATCTGTG---CAA 1034
 DB 620 CATGAGCAGAGCAACATGTCTGTGCGCCATAAACCTGTCTGGAACGTCTTGTGCGCGA 679
 QY 1035 AATATTCTACACTGGAGCCAGG-CAGTCAATTTCTGGTG-AACTCCGGGCATTTGACTGGG 1092
 DB 680 ACCATTCTCACTGGAGCCAGCGGTGAATTTCTGGGGCACCTTCTGTCGCTTGTGACTGTG 739
 QY 1093 GGAGTGAGACCAAAATC 1110
 DB 740 GGCAAGCGCAACCAACG 757
 RESULT 9
 BF143546 669 bp mRNA linear EST 24-OCT-2000
 LOCUS 601790896F1 NCI_CGAP_Lu30 Mus musculus cDNA clone IMAGE:4021870 5',
 DEFINITION mRNA sequence.
 ACCESSION BF143546
 VERSION BF143546
 KEYWORDS EST.
 SOURCE Mus musculus (house mouse)
 ORGANISM Mus musculus
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
 REFERENCE 1 (bases 1 to 669)
 AUTHORS NIH-MGC <http://mgc.nci.nih.gov/>.
 TITLE National Institutes of Health, Mammalian Gene Collection (MGC)
 JOURNAL Unpublished
 COMMENT Contact: Robert Strausberg, Ph.D.
 Email: cgapbs-remail.nih.gov
 Tissue Procurement: Gilbert Smith, Ph.D.
 cDNA Library Preparation: Life Technologies, Inc.
 cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)
 DNA Sequencing by: Incyte Genomics, Inc.
 Clone distribution: NCI-CGAP clone distribution information can be found through the I.M.A.G.E. Consortium/LLNL at:
<http://image.llnl.gov>
 Plate: LHAM9277 row: 0 column: 23
 High quality sequence stop: 659.
 Location/Qualifiers
 1. 669
 /organism="Mus musculus"
 /mol_type="mRNA"
 /strain="CZECH II"
 /db_xref="taxon:10090"
 /clone="IMAGE:4021870"
 /tissue_type="tumor, metastatic to mammary"
 /lab_host="DH10B"
 /clone_lib="NCI CGAP Lu30"
 /note="Organ: lung; Vector: pCMV-SPORT6; Site 1: NotI; Site 2: SalI; transgenic model WNT-1, expression driven by MMV-LTR enhancer; Cloned unidirectionally. Primer: Oligo dt. Library constructed by Life Technologies. Investigator providing samples: Gilbert Smith, NIH"
 BASE COUNT 195 a 140 c 165 g 169 t
 ORIGIN
 Query Match 24.1%; Score 492; DB 10; Length 669;
 Best Local Similarity 85.5%; Pred. No. 1.2e-69;
 Matches 571; Conservative 0; Mismatches 99; Indels 2; Gaps 2;
 QY 82 TAGTTGTTACATTTGGCAGGAA-AAAAATAATGCAAGATTTGCAATTTGGAACCTTG 140
 DB 3 TAGTTGTTACATCTCTTAGCAACAAAGCAATGCGCATTTGACCATCTCAGAAATCTTG 62
 QY 141 TCAACAGAGTGTGTTCTCACAAGATGGAATGTCCTTCTGATTTCTGTCGCTAT 200
 DB 63 TCAAGAGTGTGACTGTTTTCACACAGATGAGATATGGCTTCTGATTTCTGTAGCATAT 122

201	Qy	ATGTTCCAGAGAAATGTGAATTCAGTACATATGCCAACTAAAGCTGTGGACCCAGAGAACA	260
123	Db	TTACTCCAAGAAATGTGAATCTCGGGAATTTGGCCACGAAAGCTGCGATCCAGAGAACA	182
261	Qy	TTCATGAATATTAGTGAATCATCAACATCAAGGCTATCCCTGTGGAGAAATATGAATC	320
183	Db	TTCATGAATATTAGGGAATCATCAACATCAAGGGTTATCCCATGAGGAGTATGAAGTT	242
321	Qy	GCAACTGGAAGATGGGTATATCTTTCTGTTAAACAGGATTCCTCGAGGCCCTAGTGCAACCT	380
243	Db	GCAACAGAGATGGGTACATCTTTCTGTGACAGAAATCCTCGGGGACAGACCGGTTA	302
381	Qy	AGAAGACAGAGTTCCAGGCGTGTGGTGTATCTGCAGCATGGCCTAGTTGGAGTGTCTAGC	440
303	Db	AAGAAGGAAGGATCCAGGCCAGTGTGTATCTGCACCATGGTCTTCTTGGGATGTCTAGC	362
441	Qy	AACTGGATTTCCAACTCTGCCAAACAAATAGCCTGGGCTTCATTTCTGGCAGATGCTGGTTT	500
363	Db	AACTGGATTTCCAACTCTGCCAAACAAAGCTGGGTTTATTTCTGGCAGATGCAAGG-TTT	421
501	Qy	GACGTGTGATGGGAAACAGACAGGGGAAAACGCTGGTCTCGAAAAACAAAGACACTCTCC	560
422	Db	GATGTGTGATGGGAAAACAGCAGAGGAAAACCTGGTCTCGGAAGACAAAGCCCTCTCC	481
561	Qy	ATAGACCAAGATGAGTTCTGGGCTTTCAGTTTATGATGAGATGCTAGGTTTGACCTTCT	620
482	Db	ATAGATCAAGATGAGTTCTGGGCTTTCAGTTTATGATGAAATGGCTAGGTTTGACCTTCCA	541
621	Qy	GCAGTGATAACTTTTATTTTCAGAAAACGGGCCAGGAAAAGATCTTATATGTCGGCTAT	680
542	Db	GCTGTGATAAACTTTTATCTTACAGAAAACGGGCCAGGAAAAGGTCATTATGTGCGGTAC	601
681	Qy	TCACAGGGACCAACCATGGGCTTTATTGCAATTTTCCACCATGCCAGAGTGGCTCAGAAA	740
602	Db	TCACAGGGACCAACCATGGGCTTTATTGCAATTTTCCAAATGCCAGAGTACTGCTCAATAC	661
741	Qy	ATCAAAAT	748
662	Db	ATCCAAAT	669

RESULT 10
AK088659
LOCUS
DEFINITION
AK088659 2138 bp mRNA linear HTC 05-DEC-2002
Mus musculus 2 days neonate thymus thymic cells cDNA, RIKEN
full-length enriched library, clone:Z43003A04 product:lysosomal
acid lipase 1, full insert sequence.
ACCESSION
VERSION
KEYWORDS
SOURCE
AK088659.1 GI:26353707
HTC; CAP trapper.
Mus musculus (house mouse)
Mus musculus
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Mus;

TITLE	JOURNAL	MEDLINE	PUBMED	REFERENCE	AUTHORS
sequencing pipeline with 384 multicapillary sequencer	Genome Res.	10 (11), 1757-1771	(2000)	20530913	Fujiwaka, S., Inoue, K., Togawa, Y., Izawa, M., Ohara, E., Watahiki, M., Yoneda, Y., Ishikawa, T., Ozawa, K., Tanaka, T., Matsuura, S., Kawai, J., Okazaki, Y., Muramatsu, M., Inoue, Y., Kira, A. and Hayashizaki, Y.
				11050861	RIKEN integrated sequence analysis (RISA) system-384-format
				4	
					Kawai, J., Shinagawa, A., Shibata, K., Yoshino, M., Itoh, M., Ishii, Y., Arakawa, T., Hara, A., Fukunishi, Y., Konno, H., Adachi, J., Fukuda, S., Aizawa, K., Izawa, M., Nishi, K., Kiyosawa, H., Kondo, S., Yamanaka, I., Saito, T., Okazaki, Y., Gojobori, T., Bono, H., Kasukawa, T., Saito, R., Kadota, K., Matsuda, H., Ashburner, M., Batalov, S., Casavant, T., Fleischmann, W., Gaasterland, T., Gless, C., King, B., Kochiwa, H., Kuehl, P., Lewis, S., Matsuo, Y., Nikaudo, I., Pesole, G., Quackenbush, J., Schriml, L. M., Staubli, F., Suzuki, R., Tomita, M., Wagner, I., Washio, T., Sakai, K., Okido, T., Furuno, M., Aono, H., Baldarelli, R., Barsh, G., Blake, J., Boffelli, D., Bojunga, N., Carninci, P., de Bonaldo, M. F., Brownstein, M. J., Bult, C., Fletcher, C., Fujita, M., Gariboldi, M., Gustincich, S., Hill, D., Hofmann, M., Hume, D. A., Kamiya, M., Lee, N. H., Lyons, P., Marchionni, L., Mashima, J., Mazzarelli, J., Mombaerts, P., Nordone, P., Ring, B., Ringwald, M., Rodriguez, I., Sakamoto, N., Sasaki, H., Sato, K., Schonbach, C., Seya, T., Shibata, Y., Storch, K. F., Suzuki, H., Togo, Oka, K., Wang, K. H., Weitz, C., Whittaker, C. C., Wilming, L., Wynshaw-Boris, A., Yoshida, K., Hasegawa, Y., Kawaji, H., Kohtsuki, S. and Hayashizaki, Y.

TITLE	Direct Submission
JOURNAL	Submitted (16-APR-2002) Yoshihide Hayashizaki, The Institute of Physical and Chemical Research (RIKEN), Laboratory for Genome Exploration Research Group, RIKEN Genomic Sciences Center (GSC), RIKEN Yokohama Institute, 1-7-22 Suehiro-cho, Tsurumi-ku, Yokohama, Kanagawa 230-0045, Japan. (E-mail: genome-resesgsc.riken.go.jp, URL:http://genome.gsc.riken.go.jp/, Tel:81-45-503-9222, Fax:81-45-503-9216)
COMMENT	cDNA library was prepared and sequenced in Mouse Genome Encyclopedia Project of Genome Exploration Research Group in Riken Genomic Sciences Center and Genome Science Laboratory in RIKEN. Division of Experimental Animal Research in Riken contributed to prepare mouse tissues.
FEATURES	Tissues were provided by Dr. John Todd (Dept. of Medical Genetics Wellcome Trust Centre for Molecular Mechanisms in Disease Wellcome Trust/MRC building Addenbrookes Hospital Cambridge) whose assistance we gratefully acknowledge.
source	Please visit our web site for further details. URL:http://genome.gsc.riken.go.jp/ URL:http://fantom.gsc.riken.go.jp/ Location/Qualifiers 1. .2138

```

/organism="Mus musculus"
/mol_type="rRNA"
/strain="NOD"
/db_xref="FANTOM DB:E430023A04"
/db_xref="taxon:10090"
/clone="E430023A04"
/cell_type="thymic cells"
/tissue_type="thymus"
/clone_lib="RIKEN full-length enriched mouse cDNA library"
/dev_stage="2 days neonate"
66..1259
CDS
/note="unnamed protein product; lysosomal acid lipase 1
(MGD) [MG: 96789, GB [NM_021460, evidence: BLASTN, 99%,
match=2084]
putative"
/codon_start=1
/protein_id="BAC40484.1"
/db_xref="GI:26353708"
/translation="MQLGLVVFVFTIGILLRVPVTVSVDVPEVNNVTEIMRWGY
PGEHSVLTGPGYILSHIRIPGRKXNHGKPRVYVYLOHGLLADSNWVTEINDSSL
GFLADAGFDVWNGSRNTWLSKHEKTLVSQDFWAFSEFEMAKYDLPASINYLK
TGOEIVYVHSGQCTIGFIAPSOPELAKIKMFLVLAFLVSLNPFASGPIQLGRUP
DPLLKMGQKQFLPOSAMKLSIHVCTHVKELCANVFLICGNEKNLNMNRVD
VYTHCPAGTSVQNNLHWGVKVKLQAPDMGSEKXNVEHYNGSPSPSYNIKNMLP
TALWGGDRDLADINDITILLTQIPKLVHKNIPEDWHLDFINGLDAPKLYDEIISL
MKYQ"
BASE COUNT 527 a 485 c 508 g 618 t
ORIGIN

```

```

Query Match 23.3%; Score 475.4; DB 11; Length 2138;
Best Local Similarity 64.1%; Pred. No. 3.6e-67;
Matches 716; Conservative 0; Mismatches 401; Indels 0; Gaps 0;
QY 242 AGCTGTGGACCGAGAGCATTCATGAATATTAGTGAATATCATCAATCAAGGCTATCC 301
DB 140 AGCTGTGGATCTGAATCAACATGAATGTGACCGAGATATCATGCGCTGGGGATATCC 199
QY 302 CTGTGAGGAATATGAATCGCACTGAAGATGGGTATATCTTCTGTATACAGATTC 361
DB 200 TGGTAGGAGACTCGGTCTGACAGGAGATGGCTATATTTCTGATTTCCCAATCCC 259
QY 362 TCGAGGCTPAGTGCACTTCAAGACAGAGTTCAGGCTGTGGTGTATCTGACAGATGG 421
DB 260 TCGTGGCGGAAGAACATTTTGGCAAGGTCCTCCAGACCATGTGTGTATCTTCAGCAGG 319
QY 422 CCTAGTTGAGGTGTAGCACTGGATTTCCAACTGCGCCCAACATAGCTGGGCTTCAT 481
DB 320 CTGTCTGGCAGATCTTAGTAATCTGGGTCAAAACATTTGACACAGCAGCTTGGGCTTCCT 379
QY 482 TCTGGCAGATGCTGTTTGTAGCTGTGGATGGGGAACAGCAGGGAACGGCTGTCTCG 541
DB 380 CTTGGCGGATGCTGGCTTTGTATGTGGATGGGAACAGCAGAGGAACACCTGTCTCT 439
QY 542 ARAACACAGACACTCTCCATAGACACAGATGAGTTCTGGCTTTCAGTTATGATGAGAT 601
DB 440 GAAGCAACAAGACTCTTTCAGTTCTCAGGATGAGTTCTGGGCTTTCAGTTTGTATGAGAT 499
QY 602 GGCTAGTTGACCTTCTGACAGTATTAATTTTTCAGAAAAACGGCCAGGAAAA 661
DB 500 GGCMAAATATGACCTACCTGCTTCAATTAATTAATTTGAATAAATGCGCCAGAAC 559
QY 662 GATCTATTATGTCGGTATTCACAGGCAACCAATGGGCTTATGCAATTTTCACCAT 721
DB 560 AATCTATTATGTCGGTCTTCTCAAGGCTGACCAATAGTTTTCATAGCTTTTTCACAGAT 619
QY 722 GCCAGAGCTGGCTCAGAAATCAAAATGATTTTTCCTTTAGCACCACATAGCAGCTGTAA 781
DB 620 GCTGAGTTGGCAAGAGATTTAAATGTTTATGCTTGGCTCCGCTGTTGCTCTCA 679
QY 782 GCATGCAAAAACCCCGGACCAAAATTTTGTGTGTCGCAATATGATGATCAAGGGAT 841
DB 680 TTTTGCTTCAGGCGGCTACTTCAATTTGGAGGCTTGGCAGACCCCTCTTCTCAAGGACAT 739

```

```

QY 842 GTTTGGCAAAAAGAAATTTCTGTATCAGACAGATTTCTCAGACAACTTGTATTACT 901
DB 740 GTTTGGACAGAGCAATTTCTCCCTCAGAGTCGATGTGAGTGGCTGAGCAATTCAGT 799
QY 902 TTGTGGCCAGGTGATTTCTGATCAGATTTGTATAGTATATATCATTTCTTGGGGTGGATT 961
DB 800 TTGACCCACAGTCATCATGAGGAGCTTTGCGCAACGCTTTCTTCTTCTGATGTGGCTT 859
QY 962 CAACACCAACAATATGAACATCAGCGAGCAGTGTATATGCTGCGCCACACTCTTGTCTGG 1021
DB 860 CAATGAGAGAAATTTAATATATGTCTAGAGTGTATATCAACACACTGCCCTGCGGG 919
QY 1022 AACATCTGTGCAAAATTTCTACACTGGAGCCAGCAGTGAATTTCTGFGAAGTCTCGGGC 1081
DB 920 AACTTCTGTGCAAAACATGTTACACTGGGGTCAAGCTTTTAAATACCGTAAAGCTTCAAGC 979
QY 1082 ATTTGACTGGGGAGTGTGAGACCAAAATCTGGAATAATGCAATCAGCAACTCCGTGAAG 1141
DB 980 TTTTGTACTGGGAGCAGTGGAAGAAGATTTATTTTCAATACACAGAGTTTCTTCCCTC 1039
QY 1142 GTACAGAGTCAGAGATATCAGCGTCCCTACAGCAATGTGGACAGGAGTTCAGGACTGGCT 1201
DB 1040 GTACACATAAAAAACATCGCTCTACCCACTGCTTTGTGGAGCGGGCGGAGTGGCT 1099
QY 1202 TTCAATCCAGAGAGCTGAAATGCTCTCTCTGAGGTGACCAACCTCATCTACCATAA 1261
DB 1100 GGCAGACATCAATGATCAGTATCTTACTGACTCAAAATCCCAAGTTGGTGTATCACA 1159
QY 1262 GAATATTTCTGAATGGGCTCAGTGGATTTCACTCTGGGGTTTGGATGCTCTCCTCAGCGTAT 1321
DB 1160 GAATTTCTGAGTGGGACCATCTGGATTTTATCTGGGTTTGGATGCTCCCTCGGAACT 1219
QY 1322 GTACATGAATCATCATCTGATGACGAGCAGGAGGAG 1358
DB 1220 GTATGATGAATTAATCAGTCTAATGAAGAATAATACCAG 1256

```

RESULT 11

```

LOCUS AK010093
DEFINITION Mus musculus adult male tongue cDNA, RIKEN full-length enriched
library, clone:2310067K20 product:TRACVLYGLYCEROL LIPASE, LINGUAL
PRECURSOR (EC 3.1.1.3) (LINGUAL LIPASE) homolog [Rattus
norvegicus], full insert sequence.
ACCESSION AK010093
VERSION AK010093.1 GI:12845297
KEYWORDS HIC; CAP trapper.
SOURCE Mus musculus (house mouse)
ORGANISM Mus musculus
REFERENCE 1
AUTHORS Carninci, P. and Hayashizaki, Y.
TITLE High-efficiency full-length cDNA cloning
JOURNAL Meth. Enzymol. 303, 19-44 (1999)
MEDLINE 99279253
PUBMED 10349636
REFERENCE 2
AUTHORS Carninci, P., Shibata, Y., Hayatsu, N., Sugahara, Y., Shibata, K.,
Itoh, M., Konno, H., Okazaki, Y., Muramatsu, M. and Hayashizaki, Y.
TITLE Normalization and subtraction of cap-trapper-selected cDNAs to
prepare full-length cDNA libraries for rapid discovery of new genes
JOURNAL Genome Res. 10 (10), 1617-1630 (2000)
MEDLINE 20499374
PUBMED 11042159
REFERENCE 3
AUTHORS Shibata, K., Itoh, M., Aizawa, K., Nagaoka, S., Sasaki, N., Carninci, P.,
Konno, H., Akizawa, J., Nishi, K., Katsunai, T., Tashiro, H., Itoh, M.,
Sumi, N., Ishii, Y., Nakamura, S., Hazama, M., Nishine, T., Harada, A.,
Yamanoto, R., Matsumoto, H., Sakaguchi, S., Ikegami, T., Kashiwagi, K.,
Fujiwara, S., Inoue, K., Togawa, Y., Izawa, M., Ohara, E., Watahiki, M.,
Yoneda, Y., Ishikawa, T., Ozawa, K., Tanaka, T., Matsuura, S., Kawai, J.,
Okazaki, Y., Muramatsu, M., Inoue, Y., Kira, A. and Hayashizaki, Y.

```


Db 672 AAGTCCCTTTAAAAAGATTTCACCTTATTCCTAAGTTTCTTCTCAAGTGATTTGGTA 731

Qy 850 AAAAGAAATTTCTGTATCAGACAGATTTCTCAGACAACTTGTATTTACCTTTTGGCC 909

Db 732 ACAAAATGTTTCATGCCCAACACTACTTAGATCAATTTCTTGCTACGGAAGTGTCTCAC 791

Qy 910 AGGTGATTTCTGTATCAGATTTTGTAGTAATATCATGTATTTCTTGGTGGATTTCAACACA 969

Db 792 GGGAGCTGTAGATCTTCTCTGAGCAAGCTTTATTCATCTCTGTGATTTGACAGA 851

Qy 970 ACAATATGAACTGAGCGGAGCAAGTGTATGCTGCCACACTCTTGTGTGGAACATCTG 1029

Db 852 AAAAATTAATGTGAGTGGCTTTTGTATCTGATCTGATCTAGGCGATATCCAGCAGGAACATCTA 911

Qy 1030 TGCAAAATATTTACACTGAGCGAGCGAGTGAATTTCTGTGAGTCACTCCGGGATTTGACT 1089

Db 912 CTCAAGACCTTTTCCACTGGGCAAGCTTGTCTAAATCTGGAGCTTCAAGCTTAACT 971

Qy 1090 GGGGAGTGAGACCAAAAATCTCGAAAATGCAATCAGCCAACTCTCTGTAAGGTACAGAG 1149

Db 972 GGGGAGTCTTACAGAACATGTTACACTACAAATCAGAAAACGCTCTCTACTATGATG 1031

Qy 1150 TCAGAGATGACGCTCCCTACAGCATGTGGACAGAGGTGACGACTGGCTTCAATTC 1209

Db 1032 TGTGAGCCATGACCGTGGCAATTTGCACTGAGGACGCTGGCCATGACATCTGCTGATC 1091

Qy 1210 CAGAAGACGTGAAATGCTCTCTCTGAGGTGACCAACCTCATCTACCAATGAATATTC 1269

Db 1092 CCCAAGATGTCGAATCTGCTTCCCAAACTCCCAACCTTCTGTACCAATGAGGATTC 1151

Qy 1270 CTGATGGCTCAGTGATGATTCATCTGGGGTTTGGATGCTCTCCAGTATGATCAAGT 1329

Db 1152 TTCCCTACAAATCAGCTGATCTTATCTGCGGATGATGCGCTCAAGAGTTTACAATG 1211

Qy 1330 AAATCATCCATCTGATG 1346

Db 1212 AGATAGTACCATGATG 1228

RESULT 12

AK010116

LOCUS

DEFINITION Mus musculus adult male tongue cDNA, RIKEN full-length enriched library, clone:2310068005 product:TRIACTYLGLYCEROL LIPASE, LINGUAL PRECURSOR (EC 3.1.1.3) (LINGUAL LIPASE) homolog [Rattus norvegicus], full insert sequence.

ACCESSION AK010116

VERSION AK010116.1 GI:12845334

KEYWORDS HTC; CAP trapper.

SOURCE Mus musculus (house mouse)

ORGANISM Mus musculus

REFERENCE 1 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.

2 Carninci, P. and Hayashizaki, Y. High-efficiency full-length cDNA cloning Meth. Enzymol. 303, 19-44 (1999)

3 99279253 PUBLISHED 10349636

REFERENCE 2

AUTHORS Carninci, P., Shibata, Y., Hayatsu, N., Sugahara, Y., Shibata, K., Itoh, M., Kono, H., Okazaki, Y., Muramatsu, M. and Hayashizaki, Y.

TITLE Normalization and subtraction of cap-trapper-selected cDNAs to prepare full-length cDNA libraries for rapid discovery of new genes

JOURNAL Genome Res. 10 (10), 1617-1630 (2000)

MEDLINE 20499374

PUBLISHED 11042159

REFERENCE 3

AUTHORS Shibata, K., Itoh, M., Aizawa, K., Nagaoka, S., Sasaki, N., Carninci, P., Kono, H., Akiyama, J., Nishi, K., Kiteunai, T., Tashiro, H., Itoh, M., Suni, N., Ishii, Y., Nakamura, S., Hazama, M., Nishine, T., Harada, A., Yamamoto, R., Matsumoto, H., Sakaguchi, S., Ikegami, T., Kashiwagi, K., Fujiwaka, S., Inoue, K., Togawa, Y., Izawa, M., Ohara, E., Watahiki, M.,

Yoneda, Y., Ishikawa, T., Ozawa, K., Tanaka, T., Matsuura, S., Kawai, J., Okazaki, Y., Muramatsu, M., Inoue, Y., Kira, A. and Hayashizaki, Y. RIKEN integrated sequence analysis (RISA) system-384-format sequencing pipeline with 384 multicapillary sequencer Genome Res. 10 (11), 1757-1771 (2000)

20530913

11076861

REFERENCE AUTHORS

4 Kawai, J., Shinagawa, A., Shibata, K., Yoshino, M., Itoh, M., Ishii, Y., Arakawa, T., Hara, A., Fukunishi, Y., Konno, H., Adachi, J., Fukuda, S., Aizawa, K., Izawa, M., Nishi, K., Kiyosawa, H., Kondo, S., Yamanaka, I., Saito, T., Okazaki, Y., Gojobori, T., Bono, H., Kasukawa, T., Saito, R., Kadota, K., Matsuoka, H., Ashburner, M., Batalov, S., Casavant, T., Fleischmann, W., Gaasterland, T., Gissi, C., King, B., Kochiwa, H., Kuehl, P., Lewis, S., Matsuo, Y., Nikaide, I., Pesole, G., Quackenbush, J., Schriml, L. M., Stabli, F., Suzuki, R., Tomita, M., Wagner, L., Washio, T., Sakai, K., Okido, T., Furuno, M., Aono, H., Baldairelli, R., Barsh, G., Blake, J., Boffelli, D., Bojunga, N., Carninci, P., de Bonaldo, M. F., Brownstein, M. J., Bult, C., Fletcher, C., Fujita, M., Gariboldi, M., Gustincich, S., Hill, D., Hofmann, M., Hume, D. A., Kamiya, M., Lee, N. H., Lyons, P., Marchionni, L., Mashima, J., Mazzarelli, J., Mombaerts, P., Nordone, P., Ring, B., Ringwald, M., Rodriguez, I., Sakamoto, N., Sasaki, H., Sato, K., Schonbach, C., Seya, T., Shibata, Y., Storch, K. F., Suzuki, H., Toyooka, K., Wang, K. H., Weitz, C., Whitaker, C., Wilming, L., Wynshaw-Boris, A., Yoshida, K., Hasegawa, Y., Kawaji, H., Kohtsuki, S. and Hayashizaki, Y.

Functional annotation of a full-length mouse cDNA collection Nature 409 (6821), 685-690 (2001)

21085660

11217851

REFERENCE AUTHORS

5 The PANTOM Consortium and the RIKEN Genome Exploration Research Group Phase I & II Team. Analysis of the mouse transcriptome based on functional annotation of 60,770 full-length cDNAs Nature 420, 563-573 (2002)

6 (bases 1 to 1347)

Adachi, J., Aizawa, K., Akahira, S., Akimura, T., Arai, A., Aono, H., Arakawa, T., Bono, H., Carninci, P., Fukuda, S., Fukunishi, Y., Furuno, M., Hanagaki, T., Hara, A., Hayatsu, N., Hiramoto, K., Hirakawa, T., Hori, F., Imotani, K., Ishii, Y., Itoh, M., Izawa, M., Kasukawa, T., Kawai, J., Kojima, Y., Kono, H., Konda, M., Koya, S., Kurihara, C., Matsuyama, T., Miyazaki, A., Nishi, K., Nomura, K., Numazaki, R., Ohno, M., Okazaki, Y., Okido, T., Owa, C., Saito, H., Saito, R., Sakai, C., Sakai, K., Sano, H., Sasaki, D., Shibata, K., Shibata, Y., Shinagawa, A., Shiraki, T., Sogabe, Y., Suzuki, K., Tashiro, M., Tagawa, A., Takahashi, F., Tanaka, T., Tejima, Y., Toyota, T., Yamamura, T., Yasunishi, A., Yoshida, K., Yoshino, M., Muramatsu, M. and Hayashizaki, Y.

Direct Submission

REFERENCE AUTHORS

Submitted (10-JUL-2000) Yoshihide Hayashizaki, The Institute of Physical and Chemical Research (RIKEN), Laboratory for Genome Exploration Research Group, RIKEN Genomic Sciences Center (GSC), RIKEN Yokohama Institute, 1-7-22 Suehiro-cho, Tsurumi-ku, Yokohama, Kanagawa 230-0045, Japan [E-mail: genome-res@sc.riken.go.jp, URL: http://genome.gsc.riken.go.jp/, Tel: 81-45-503-9222, Fax: 81-45-503-9216]

Please visit our web site (<http://genome.gsc.riken.go.jp/>) for further details.

REFERENCE AUTHORS

CDNA library was prepared and sequenced in Mouse Genome Encyclopedia Project of Genome Exploration Research Group in Riken Genomic Sciences Center and Genome Science Laboratory in RIKEN. Division of Experimental Animal Research in Riken contributed to prepare mouse tissues. First strand cDNA was primed with a primer [5'-GAGAGAGAGAGGATCAAGCTCTTTTCTTTTNN 3'], cDNA was prepared by using trehalose thermo-activated reverse transcriptase and subsequently enriched for full-length by cap-trapper. Second strand cDNA was prepared with the primer adapter of sequence [5'-GAGAGAGAGATCTGCTGATTAATTAATATCCCCCCCCC 3']. cDNA was cleaved with XhoI and SgrI. Cloning sites, 5' end: XhoI; 3' end: SgrI. Host: SOLR.

REFERENCE AUTHORS

Submitted (10-JUL-2000) Yoshihide Hayashizaki, The Institute of Physical and Chemical Research (RIKEN), Laboratory for Genome Exploration Research Group, RIKEN Genomic Sciences Center (GSC), RIKEN Yokohama Institute, 1-7-22 Suehiro-cho, Tsurumi-ku, Yokohama, Kanagawa 230-0045, Japan [E-mail: genome-res@sc.riken.go.jp, URL: http://genome.gsc.riken.go.jp/, Tel: 81-45-503-9222, Fax: 81-45-503-9216]

Please visit our web site (<http://genome.gsc.riken.go.jp/>) for further details.

REFERENCE AUTHORS

CDNA library was prepared and sequenced in Mouse Genome Encyclopedia Project of Genome Exploration Research Group in Riken Genomic Sciences Center and Genome Science Laboratory in RIKEN. Division of Experimental Animal Research in Riken contributed to prepare mouse tissues. First strand cDNA was primed with a primer [5'-GAGAGAGAGAGGATCAAGCTCTTTTCTTTTNN 3'], cDNA was prepared by using trehalose thermo-activated reverse transcriptase and subsequently enriched for full-length by cap-trapper. Second strand cDNA was prepared with the primer adapter of sequence [5'-GAGAGAGAGATCTGCTGATTAATTAATATCCCCCCCCC 3']. cDNA was cleaved with XhoI and SgrI. Cloning sites, 5' end: XhoI; 3' end: SgrI. Host: SOLR.

FEATURES

Location/Qualifiers

Db 612 TGGCTAAAAAATCAGAGGTTTTATGTCATTTAGTCCAGTCTGCTACTGTGAAGTATACAG 671
 QY 790 AAGCCCGGGACCAAAATTTTGTGCTCCAGATATGATGATCAAGGATTTTGGCA 849
 Db 672 AAGTCCCTTTAAAAAGATTTCACTATTTCCTTAAGTTCTTCTCAAGGTGATATTGGTA 731
 QY 850 AAAAAGAAATTTCTGTATCAGACAGATTTCTCAGACAACTTGTATTATTTACCTTTGTGGCC 909
 Db 732 ACAAAATGTTTATGCCCCACAACTACTTAGATCAATTTCTTGGTAGGAAGTGTCTCAC 791
 QY 910 AGGTGATTTCTGATCAGATTTGTAGTAATATCATGTTACTTCTGGTGGATTCACACCA 969
 Db 792 GGAGCTGTGATGATCTTCTGACAGCAAGCTTTATTCATCTCTGTGGAATTTGACAGA 851
 QY 970 ACAATATGAATGAGCCGAGCAAGTGTATGCTGCGCCACACTCTTGTGGAACATCTG 1029
 Db 852 AAAACTTAAATGTGAGTCTGTTGGTGTATCTAGGCGATATCCAGCAGGAACATCTA 911
 QY 1030 TGCATAATTTCTACTGAGCCAGCAGTGAATTTCTGTTGAACTCTCGGGCATTTGACT 1089
 Db 912 CTCAGACCTTTTCCACTGGCAGCAGCTTGTCTAAATCTGGGAAGCTTCAAGCCTATACT 971
 QY 1090 GGGGAGTGAAGCAAAATCTGAAAAATGCAATCAGCAACTCTCTGTAAGGTACAGAG 1149
 Db 972 GGGGAAGTCCATACAGACATGTTACACTACATCAGAAAAGCTCCCTACTATGATG 1031
 QY 1150 TCAGATATGACGGTCTCTACAGCAATGTGACAGAGGTGAGGATGCTGCTTCAATC 1209
 Db 1032 TGTGAGCCATGACCGTGGCAATGTCAGTGTGGAACGGTGGCCATGACATCTCGGTGATC 1091
 QY 1210 CAGACAGCTGAAATGCTCTCTGAGGTGACCAACTCATCTACCATAGATATTC 1269
 Db 1092 CCCAGATGTGCAATGCTGTTCCAACTCCCACTCTGTTACCATAGGAGATTC 1151
 QY 1270 CTGAATGGGCTACGTTGATTTCAATCGGGTTGGATGCTCTCCCTGATGTCATATG 1329
 Db 1152 TTCCCTACATCACTACCTGGACTTCATCTGGCGATGATCGGCTCAAGAGGTTTCAATG 1211
 QY 1330 AAATCATCCATCTGATG 1346
 Db 1212 AGATAGTTACCATGATG 1228

 RESULT 14
 AK010139
 LOCUS
 DEFINITION Mus musculus adult male tongue cDNA, RIKEN full-length enriched library, clone:2310069P19 product:TRACYLGLYCEROL LIPASE, LINGUAL PRECURSOR (BC 3.1.1.3) (LINGUAL LIPASE) homolog [Rattus norvegicus], full insert sequence.
 ACCESSION AK010139
 VERSION 1 GI:12845371
 KEYWORDS HTC; CAP trapper.
 SOURCE Mus musculus
 ORGANISM Mus musculus (house mouse)
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
 REFERENCE 1
 AUTHORS Carninci, P., Shibata, Y., Hayatsu, N., Sugahara, Y., Shibata, K., Itoh, M., Konno, H., Okazaki, Y., Muramatsu, M. and Hayashizaki, Y.
 TITLE Normalization and subtraction of cap-trapper-selected cDNAs to prepare full-length cDNA libraries for rapid discovery of new genes
 JOURNAL Genome Res. 10 (10), 1617-1630 (2000)
 MEDLINE 20493374
 PUBMED 11042159
 REFERENCE 2
 AUTHORS Carninci, P., Shibata, Y., Hayatsu, N., Sugahara, Y., Shibata, K., Itoh, M., Konno, H., Okazaki, Y., Muramatsu, M. and Hayashizaki, Y.
 TITLE Normalization and subtraction of cap-trapper-selected cDNAs to prepare full-length cDNA libraries for rapid discovery of new genes
 JOURNAL Genome Res. 10 (10), 1617-1630 (2000)
 MEDLINE 20493374
 PUBMED 11042159
 REFERENCE 3
 AUTHORS Shibata, K., Itoh, M., Aizawa, K., Nagaoka, S., Sasaki, N., Carninci, P.,

Konno, H., Akiyama, J., Nishi, K., Kitsumai, T., Tashiro, H., Itoh, M., Suni, N., Ishii, Y., Nakamura, S., Hazama, M., Nishine, T., Harada, A., Yamamoto, R., Matsumoto, H., Sakaguchi, S., Ikegami, T., Kashiwagi, K., Fujiwaka, S., Inoue, K., Togawa, Y., Izawa, M., Ohara, E., Watanabe, M., Yoneda, Y., Ishikawa, T., Ozawa, K., Tanaka, T., Matsura, S., Kawai, J., Okazaki, Y., Muramatsu, M., Inoue, Y., Kira, A. and Hayashizaki, Y.
 RIKEN integrated sequence analysis (RISA) system--384-format sequencing pipeline with 384 multicapillary sequencer
 Genome Res. 10 (11), 1757-1771 (2000)
 20530913
 11076861
 4

Kawai, J., Shinagawa, A., Shibata, K., Yoshino, M., Itoh, M., Ishii, Y., Arakawa, K., Hara, A., Fukunishi, Y., Konno, H., Adachi, J., Fukuda, S., Aizawa, K., Izawa, M., Nishi, K., Kiyosawa, H., Kondo, S., Yamanaka, I., Saito, T., Okazaki, Y., Gojohori, T., Bono, H., Kasukawa, T., Saito, R., Kadota, K., Matsuda, H., Ashburner, M., Batalov, S., Casavant, T., Fleischmann, W., Gaasterland, T., Gissi, C., King, B., Kochiwa, H., Kuehl, P., Lewis, S., Matsuo, Y., Nikaido, I., Pesole, G., Quackenbush, J., Schriml, L. M., Stabli, F., Suzuki, R., Tomita, M., Wagner, L., Washio, T., Sakai, K., Okido, T., Furuno, M., Aono, H., Balderelli, R., Barsh, G., Blake, J., Boffelli, D., Bojunga, N., Carninci, P., de Bonaldo, M. F., Brownstein, M. J., Bult, C., Fletcher, C., Fujita, M., Gariboldi, M., Gustincich, S., Hill, D., Hofmann, T., Hume, D. A., Kamiya, M., Lee, N. H., Lyons, P., Marchionni, L., Mashima, J., Mazzarelli, J., Mombarts, P., Nordone, P., Ring, B., Ringwald, M., Rodriguez, I., Sakamoto, N., Sasaki, H., Sato, K., Schonbach, C., Seya, T., Shibata, Y., Storch, K. F., Suzuki, H., Toyooka, K., Wang, K. H., Weitz, C., Whittaker, C., Wilming, L., Wynshaw-Boris, A., Yoshida, K., Hasegawa, Y., Kawaji, H., Kohseuki, S. and Hayashizaki, Y.
 Functional annotation of a full-length mouse cDNA collection
 Nature 409 (6821), 685-690 (2001)
 21085660
 11217851

TITLE
 JOURNAL
 MEDLINE
 PUBMED
 REFERENCE
 AUTHORS

5
 The PANTOM Consortium and the RIKEN Genome Exploration Research Group Phase I & II Team.
 Analysis of the mouse transcriptome based on functional annotation of 60,770 full-length cDNAs
 Nature 420, 563-573 (2002)
 6 (bases 1 to 1350)
 Arakawa, T., Bono, H., Carninci, P., Fukuda, S., Fukunishi, Y., Furuno, M., Hanagaki, T., Hara, A., Hayatsu, N., Hiramoto, K., Hiroaka, T., Hori, F., Imotani, K., Ishii, Y., Itoh, M., Izawa, M., Kasukawa, T., Kato, H., Kawai, J., Kojima, Y., Konno, H., Kouda, M., Koya, S., Kurihara, C., Matsuyama, T., Miyazaki, A., Nishi, K., Nomura, K., Numazaki, R., Ohno, M., Okazaki, Y., Okido, T., Owa, C., Saito, H., Saito, R., Sakai, C., Sakai, K., Sano, H., Sasaki, D., Shibata, K., Shibata, Y., Shinagawa, A., Shiraki, T., Sogabe, Y., Suzuki, H., Tagami, M., Tagawa, A., Takahashi, F., Tanaka, T., Tejima, Y., Toya, T., Yamamura, T., Yasunishi, A., Yoshida, K., Yoshino, M., Muramatsu, M. and Hayashizaki, Y.
 Direct Submission
 Submitted (10-JUL-2000) Yoshihide Hayashizaki, The Institute of Physical and Chemical Research (RIKEN), Laboratory for Genome Exploration Research Group, RIKEN Genomic Sciences Center (GSC), RIKEN Yokohama Institute, 1-7-22 Suehiro-cho, Teurumi-ku, Yokohama, Kanagawa 230-0045, Japan (E-mail: genome-res@gscc.riken.go.jp, URL: http://genome.gsc.riken.go.jp/, Tel: 81-45-503-9222, Fax: 81-45-503-9216)
 Please visit our web site (http://genome.gsc.riken.go.jp/) for further details.

TITLE
 JOURNAL
 MEDLINE
 PUBMED
 REFERENCE
 AUTHORS

TITLE
 JOURNAL

CNA library was prepared and sequenced in Mouse Genome Encyclopedia Project of Genome Exploration Research Group in Riken Genomic Sciences Center and Genome Science Laboratory in Riken. Division of Experimental Animal Research in Riken contributed to prepare mouse tissues. First strand cDNA was primed with a primer [5' GAGAGAGAGAGATCCAGAGCTCTTTTATTTTNN 3'], cDNA was prepared by using trihalose thermo-activated reverse transcriptase and subsequently enriched for full-length by cap-trapper. Second strand cDNA was prepared with the primer adapter of sequence [5'

COMMENT

GAGAGAGAGATTCCTGAGTAAATTAATTAATCCCTCCCTCCCTCCCT 3'. cDNA was cleaved with XhoI and SstI. Cloning sites, 5' end: XhoI; 3' end: SstI.
Host: SstI.

FEATURES

Location/Qualifiers
1. 1350
/organism="Mus musculus"
/mol_type="mRNA"
/strain="C57BL/6J"
/db_xref="FANTOM.DB:2310069P19"
/db_xref="PMI:1898377"
/db_xref="taxon:10090"
/clone="2310069P19"
/sex="male"
/tissue_type="tongue"
/clone_lib="RIKEN full-length enriched mouse cDNA library"
/dev_stage="adult"
53..1240
/note="unnamed protein product; TRIACYLGLYCEROL LIPASE, LINGUAL PRECURSOR (EC 3.1.1.3) (LINGUAL LIPASE) homolog [Rattus norvegicus] (SWISSPROT:P04634, evidence: FASTA, 88.4%ID, 100%length, match=1184) putative"
/codon_start=1
/protein_id="BAB26725.1"
/db_xref="GI:12845372"
/db_xref="PMI:1914967"
/translation="MWLLVTSVLSAFGGAHLFGKLGPKNPEANNVNSQMITYWGYP SEYVEDVIGLVNRYPIYKKNSENIQKRVAVLQHLIASATNWTINPNLSA FILADAGVYVLSNRTNRYKNNVYSPDSFVAFSPDMAYKDLPTIDIVOKT QEKIHYVNSQGTGFIAPSTNPALAKIKFVYALPIATVYKTESPKKLSLIPK FLKVIIPGNKMFPHNYLDOLFTGVCSRELLDLCSNLFIFCGPDKNLNVSREDV YLGNPAGTSDQLFWAQLAKSKLOAYNWSPLQNLHYNKQTPPYDVDSANTVPI AVRWGHDILADPDVAVLLPFLKPLNLIHKLPIHLLDFINADAPQEVINEIVTMM AED"

CDS

BASE COUNT 398 a 289 c 283 g 380 t
ORIGIN
Query Match 22.5%; Score 460.2; DB 11; Length 1350;
Best Local Similarity 63.7%; Pred. No. 1.2e-64;
Matches 599; Conservative 0; Mismatches 398; Indels 0; Gaps 0;
QY 250 ACCGAGAGCATTCATGAATATTAGTGAATCATCCCAACATCAAGGCTATCCCTGTGAGG 309
DB 132 ACCCTGAAGCAACATGAATCTTAGTCAGATGATACTTCTGGGGATATCCAAAGTGAGG 191
QY 310 AATATGAAGTGCACATGAGTGGTATATCTTCTGTAAACAGGATTCCTCGAGCC 369
DB 192 AATATGAAGTGTACTGAAGATGCTACATCTGGGGGCTATAGAAATTCCTTATGGGA 251
QY 370 TAGTGCAACCTAAGAAGACAGGTTCCAGGCTGTGGTGTACTGCAGCATGGCCTAGTTG 429
DB 252 AGAAAAATCTGAGAAATATCGCAAGAGACCTGTGGCATATTTTCAGCATGGTTTGATTG 311
QY 430 GAGGTGCTAGCAATCGGATTTCCAACTGCCCAACATAGCTGGGCTTCATCTTGGCAG 489
DB 312 CATGAGCCACAACATGGGATTAACAATCTGCCAAACACAGCCTGGCCTTCATCTAGCAG 371
QY 490 ATGCTGGTTTTCAGCTGTGGATGGGAAACAGCAGGGAACCGCTGTCTCGAAAAACACA 549
DB 372 ATGCTGGCTATGATGTGGCTGGGAACAGTCCAGGGAATACATGCTCCGGGAAATG 431
QY 550 AGACACTTCATAGACCAAGATGATGTTCTGGGCTTCAGTTATGATGAGATGGCTAGGT 609
DB 432 TATACTATTCACCAAGACTCAATTCAGGCTTTTCAGCTTTGATGAATGGCTAAAT 491
QY 610 TTGACCTTCCTGCACTGATAAATCTTATTTTCAGAAAAACGGGCGAGGAAAGATCTATT 669
DB 492 ATGACCTTCAGCCACCATAGCTTCATTTGACAGAAACTGGACAGAGAGATCACT 551
QY 670 ATGCGGCTATTCAGGGGACCAACATAGGCTTTATGCAATTTCCACCATGCCAGGC 729
DB 552 ATGTTGGTCACTCTCAGGGCCCACTATCGGTTTTATTGCGCTTTTCTACCAATCCTGCTC 611

QY 730 TGGCTCAGAAATCAAAATGATATTTGCTTTAGCACCCATAGCCACCTGTTAAGCATCAA 789
DB 612 TGGCTAAATAATCAAGAGGTTTTATGATTTAGCTCCAAATGCTGCTACTGTGAGTATACAG 671
QY 790 AAAGCCCGGAGCAAAATTTTTGTTGCTGCAGATATGATCAAGGATTTGTTGGCA 849
DB 672 AAAGTCCCTTTAAAAAGATTTCACTTATTCCTAAGTTCTTCTCAAGGTGATATTTGGTA 731
QY 850 AAAAGAAATTTCTGATCAGACCAAGATTTCTCAGACAACTGTTATTTACCTTTGTGGCC 909
DB 732 ACAAAATGTTCAATGCCCAACAACTATTAGATCAATTTCTTGGTACGGAAGTGTGCTCAC 791
QY 910 AGGTGATTTCTGATCAGATTTGTAGTAAATATCATGTTACTTTCTGGTGGATTCACACCA 969
DB 792 GGGAGCTGTAGATCTTCTGCAAGCAAGCTTTATTTATTCATCTTCTGTGGATTTGACAAGA 851
QY 970 ACATATGAACATGAGCCGAGCAAGTGTATATGCTGCCACACTCTTCTGCTGGAACATCTG 1029
DB 852 AAAACTTAAATGTGAGTGCCTTTGATGTGATCTAGGGCAATAATCCAGCAGGAACATCTA 911
QY 1030 TGCAAAATATTCTACACTGAGCCAGGCAATTTCTGTTGAATCTTCTGGTGAACCTCCGGGCA 1089
DB 912 CTCAGAGACTTTTCCACTGGGCACAGCTTGTCTAAATCTGGGAAGCTTCAAGSCTATACT 971
QY 1090 GGGGAGTGAAGCAAAATCTGAAAAATCAATCAGCAACTCCTGTAAGGTACAGAG 1149
DB 972 GGGGAAGTCCATTACAGAACTGTTACTACTACAACTGAAAAACGCTCCCTACTATGATG 1031
QY 1150 TCAGAGATATGACGGTCCCTCAGCAATGTGCAAGAGAGGTGAGGATCTGGCTTTCAAATC 1209
DB 1032 TGTGAGCCATGACCGTGCCTTTCAGTGTGGAAGGTGGGCAATGACATCTGGCTGATC 1091
QY 1210 CAGAAGAGCTGAAATATGCTCTCTGAGGTGACCAACCTCATCTACCAATGAATATTC 1269
DB 1092 CCCAAGATGTGCAATGCTGCTTCCCAACTCCCAACCTTTTGTACCATCAAGGATTC 1151
QY 1270 CTGAATGGCTCAGCTGATTTCTCTGGGTTTGGATGCTCTCACCGTATGTACATG 1329
DB 1152 TTCCCTACAACTCACTGAGCTTCTCTGGGCGATGATGCGCTTCAGAGGTTTACATG 1211
QY 1330 AATCATCATCTGTATG 1346
DB 1212 AGATAGTTACCATGATG 1228

RESULT 15
AK010124
LOCUS
DEFINITION
AK010124 1340 bp mRNA linear HTC 05-DEC-2002
Mus musculus adult male tongue cDNA, RIKEN full-length enriched library, clone:2310069P09 product:TRIACYLGLYCEROL LIPASE, LINGUAL PRECURSOR (EC 3.1.1.3) (LINGUAL LIPASE) homolog [Rattus norvegicus], full insert sequence.
ACCESSION
AK010124
VERSION
AK010124.1 GI:12845346
KEYWORDS
HTC; Cat trapper.
SOURCE
Mus musculus (house mouse)
ORGANISM
Mus musculus
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
REFERENCE
1
Carninci, P. and Hayashizaki, Y.
High-efficiency full-length cDNA cloning
Meth. Enzymol. 303, 19-44 (1999)
99279253
PUBMED 10349636
2
Carninci, P., Shibata, Y., Hayatsu, N., Sugahara, Y., Shibata, K., Itoh, M., Kono, H., Okazaki, Y., Muramatsu, M. and Hayashizaki, Y.
Normalization and subtraction of cap-trapper-selected cDNAs to prepare full-length cDNA libraries for rapid discovery of new genes
Genome Res. 10 (10), 1617-1630 (2000)
20499374
PUBMED 11042159

Db	546	ATGTTGCTCACTCTCAGGACCACTATCGGTTTATTGGCTTTTCTACCAATCCTGCTC	605
Qy	730	TGGCTCAGAAATCAAAATGTATTTTCTTAGCACCATAGCCACTGTTAAGCATGCAA	789
Db	606	TGGCTAAAAAATCAAGAGGTTTTATGCAATAGCTCCAGTTGCTAGTGAAGTATACAG	665
Qy	790	AAAGCCCGGAGCAAAATTTTGTGCTGCAGATATGATCAAGGGATTGTTGGCA	849
Db	666	AAAGTCCCTTAAAAAGATTCTACTTATCTTAAGTTTCTCTCAGGTGATATTTGGTA	725
Qy	850	AAAAAGAAATTTCTGTATCAGACCAAGATTCTCAGACAACTTGTATTATTCCTTTTGGCC	909
Db	726	ACAAAATGTTTCATGCCCACTACTAGATCAATTTCTTGGTACGGAAGTGTGCTCAC	785
Qy	910	AGTGTATCTTGATCAGATTGTAGTAATATCATGTTACTTCTGGGTGGATTCAACACCA	969
Db	786	GGGAGCTGCTAGATCTCTCTGACGCAACGCTTATTTCACTCTCTGTGGATTGACAGA	845
Qy	970	ACAATATGAACATGAGCCGAGCAAGTGTATATGCTGCCCACTCTTTGCTGGAACTCTG	1029
Db	846	AAAACTTAAATGTAGTGGCTTTTGTATCTAGGGCATAATCCAGCAGGAACATCTA	905
Qy	1030	TGCAAAATATTTCTACACTGGAGCCAGCAGTGAATTTCTGGTGAATCCGGGCAATTTGACT	1089
Db	906	CTCAGACCTTTTCCACTGGGCAAGCTTGTCTAAATCTGGAGCTTCAAGCCCTATAACT	965
Qy	1090	GGGGAGTGAGACCAAAATCTGAAAAATGCAATCAGCCCACTCTCTGTAAAGGTACAGAG	1149
Db	966	GGGGAAGTCCATTACAGAAATGTTTACACTACATCAGAAAAAGCCCTCCCTACTATGATG	1025
Qy	1150	TCAGAGATATCAGCTCCCTACAGCAATGTGRCAGGAGTCAAGGCTGGCTTTCAAATC	1209
Db	1026	TGTAGCCATGACCGTGCCAAATTCAGTGTGGAACGGTGGCCATGATCTCTGGCTGATC	1085
Qy	1210	CAGAAGCGTGAATATGCTCTCTGAGGTGACCAACCTCATCTACCATPAAGAAATTC	1269
Db	1086	CCCAAGATGTGCAATGTGCTTCCCAAACTCCCAACCTTCTGTACCATAGGAGATTC	1145
Qy	1270	CTGAATGGCTCAGCTGGATTTCATCTGGGTTTGGATGCTCTCCCTCACCCTATGTACAATG	1329
Db	1146	TCCCTACAATCACCTGGACTTTCATCTGGGCGATGGATGGCCCTCAGGAGTTTACATG	1205
Qy	1330	AAATCATCCATCTGATG	1346
Db	1206	AGATAGTTACCATGATG	1222

Search completed: February 19, 2004, 10:30:12
 Job time : 4478.52 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2004 Compugen Ltd.

OM nucleic - nucleic search, using sw model

Run on: February 19, 2004, 04:49:02 ; Search time 126.478 Seconds
(without alignments)
7133.182 Million cell updates/sec

Title: US-10-042-431-45
Perfect score: 2044
Sequence: 1 gtgacccacgctccggg.....aaaaaaaaaagggcgccgc 2044

Scoring table: IDENTITY NUC
Gapop 10.0 , Gapext 1.0

Searched: 569978 seqs, 220691566 residues

Total number of hits satisfying chosen parameters: 1139956

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%
Listing first 45 summaries

Database : Issued Patents NA.*
1: /cgm2_6/ptodata/2/ina/5A_COMB.seq.*
2: /cgm2_6/ptodata/2/ina/5B_COMB.seq.*
3: /cgm2_6/ptodata/2/ina/6A_COMB.seq.*
4: /cgm2_6/ptodata/2/ina/6B_COMB.seq.*
5: /cgm2_6/ptodata/2/ina/PTCUS_COMB.seq.*
6: /cgm2_6/ptodata/2/ina/backfiles.seq.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	470	23.0	1367	4	US-09-348-930A-8
2	456.8	22.3	1411	4	US-09-185-489-3
3	456.8	22.3	1411	4	US-10-043-665B-3
4	453.6	22.2	1134	4	US-09-185-489-4
5	453.6	22.2	1134	4	US-10-043-665B-4
6	450.4	22.0	1134	4	US-09-185-489-1
7	450.4	22.0	1134	4	US-10-043-665B-1
8	449.8	22.0	8324	4	US-09-185-489-5
9	449.8	22.0	8324	4	US-10-043-665B-5
10	444.2	21.7	1137	1	US-08-227-108-2
11	444.2	21.7	1137	2	US-09-073-674-2
12	444.2	21.7	1140	1	US-08-227-108-4
13	444.2	21.7	1140	2	US-09-073-674-4
14	444.2	21.7	1146	1	US-08-227-108-6
15	444.2	21.7	1146	2	US-09-073-674-6
16	444.2	21.7	1198	4	US-09-348-930A-5
17	444.2	21.7	1528	1	US-08-227-108-1
18	444.2	21.7	1528	2	US-09-073-674-1
19	444.2	21.7	1528	4	US-09-348-930A-1
20	444.2	21.7	1528	4	US-09-348-930A-7
21	414.2	20.3	1360	4	US-09-820-001-1
22	388.8	19.0	1048	4	US-09-348-930A-3
23	104.2	5.1	22067	1	US-09-820-001-3
24	48.2	2.4	178	1	US-08-751-782-2
25	48.2	2.4	178	2	US-08-925-171-2
26	46.4	2.3	43676	3	US-08-958-952-12
27	42	2.1	8041	1	US-08-765-081-1

C	28	42	2.1	8041	3	US-09-098-082-1	Sequence 1, Appli
C	29	42	2.1	8041	5	PCT-US95-06994-1	Sequence 1, Appli
C	30	41.4	2.0	87563	4	US-09-453-702B-57	Sequence 57, Appli
C	31	40.4	2.0	1140	2	US-08-698-805-5	Sequence 5, Appli
C	32	39.8	1.9	26000	4	US-09-843-376-10	Sequence 10, Appli
C	33	39.8	1.9	168575	4	US-09-426-290-1	Sequence 1, Appli
C	34	39.4	1.9	53332	4	US-09-801-861-3	Sequence 3, Appli
C	35	39	1.9	874	4	US-09-389-681-426	Sequence 426, App
C	36	39	1.9	874	4	US-09-620-405B-426	Sequence 426, App
C	37	39	1.9	874	4	US-09-433-826B-426	Sequence 426, App
C	38	39	1.9	874	4	US-09-604-287A-426	Sequence 426, App
C	39	38.6	1.9	2815	1	US-08-230-491A-1	Sequence 1, Appli
C	40	38.6	1.9	2815	1	US-08-619-280A-1	Sequence 1, Appli
C	41	38.6	1.9	2815	2	US-08-940-391-1	Sequence 1, Appli
C	42	38.6	1.9	1664976	4	US-08-916-421B-1	Sequence 1, Appli
C	43	38.4	1.9	2798	3	US-09-318-448-34	Sequence 34, Appli
C	44	38	1.9	215	4	US-09-702-705-499	Sequence 499, App
C	45	38	1.9	215	4	US-09-736-457-499	Sequence 499, App

ALIGNMENTS

RESULT 1

US-09-348-930A-8
Sequence 8, Application US/09348930A
Patent No. 6573431
GENERAL INFORMATION:
APPLICANT: Lence, P
APPLICANT: Gruber, V
APPLICANT: Baudino, S
APPLICANT: Merot, B
APPLICANT: Benicourt, C
APPLICANT: Cudrey, C
TITLE OF INVENTION: Recombinant Predutodenal Lipases and Polypeptides Derivatives Prod
FILE REFERENCE: 18433/2012
CURRENT APPLICATION NUMBER: US/09/348,930A
CURRENT FILING DATE: 1999-07-02
PRIOR APPLICATION NUMBER: 08/945,321
PRIOR FILING DATE: 1998-02-12
PRIOR APPLICATION NUMBER: PCT/FR96/00606
PRIOR FILING DATE: 1996-03-19
NUMBER OF SEQ ID NOS: 37
SOFTWARE: Patent in version 3.0
SEQ ID NO 8
LENGTH: 1367
TYPE: DNA
ORGANISM: Homo sapiens
FEATURE:
NAME/KEY: CDS
LOCATION: (47) .. (1240)
US-09-348-930A-8

Query Match	23.0%	Score	470	DB	4	Length	1367
Best Local Similarity	64.2%	Pred. No.	1.4e-124	Indels	0	Gaps	0
Matches	707	Conservative	0	Mismatches	395		
Qy	244	CTGTGACCCAGAACGATTCATCAATATTAGTGAATAATCATCCAAATCAAGGCTATCCCT	303				
Db	123	CTGGAAGCCCTGAAGTGACTATGAACATTAGTCAGATGATTACTTATTGGGGATACCCAA	182				
Qy	304	GTGAGAATATGAGTCGCAATCGAGATGGGTATATCTTCTTAAACAGATTCTTC	363				
Db	183	ATGAAGAATATGAAGTTGTGACTGAAGATGGTTATTCTTGAAGTCATAGATTCTTC	242				
Qy	364	GAGGCCCTAGTCACACCTAAAGAGACAGGTTCCAGGCTGTGGTGTGTACTGACGATGGCC	423				
Db	243	ATGGGAAGAAAATTCAAGGAATACAGGCAGACGCTGTGTGTGTGTGTGTGTGTGTGTGT	302				
Qy	424	TAGTGGAGGTCTAGCAACTGGATTTCACCTGCCCAACAAATAGCTGGGCTTCATTC	483				
Db	303	TGCTTGATCAGCCCAAACTGGATTTCACCTGCCCAACAACTGGCTTTCATTC	362				

484 TGGCAGATCTGGTTTGAAGTGTGGATGGGGAACAGCAGGGGAAACCGCTGCTCGAA 543
Db TGGCAGATCTGGTTTGAAGTGTGGATGGGGAACAGCAGGGGAAACCGCTGCTCGAA 422
QY 544 AACACAGACACTCTCCATAGACCAAGATGAGTCTCTGGGCTTTTCAATGATGATGAGTGG 603
Db 423 GAAACTTGTACTATTACACAGATTTCAGTTGAAATCTCTGGGCTTTTCAATGATGAAATGG 482
QY 604 CTAGTTTGAACCTCTCTGAGTGTAACTTTTATTTTGGAGAAACGGCCAGGAAAGA 663
Db 483 CTAATATGACCTTCCAGCCCAATCACTTCAATGTAAGAAAGAACTGGACAGAGCAGC 542
QY 664 TCTATTATGTGGGCTATTACAGGGCACCACATGGGCTTTATGCAATTTTCCACCATGC 723
Db 543 TACACTATGTGGCCATTCCAGGGCACCACATGGTTTTATGCTTTTCCACCAATC 602
QY 724 CAGAGCTGGCTCAGAAAATCAAAATGTATTTTGTCTTTAGCAGCCATAGCCATGTTAGC 783
Db 603 CCAGCCTGGCTAAAGAAATCAAAACCTTCTATGCTCTAGCTCTCTGTTGCACTGTGAAGT 662
QY 784 ATCAAAAAGCCCGGACCCAAATTTTGTGCTGCGCAGATATGATGATCAAGGGATTGT 843
Db 663 ATCAAAAAGCCCTTAAACAACTTAGATTTTCTCTCAATCCCTCTTCAAGTTTAT 722
QY 844 TTGGCAAAAAGAAATTTCTGTATCAGACCAAGATTTCTCAGACAACTTGTGTTATTTTACCTTT 903
Db 723 TTGGTGACAAAATATTCTACCCACACAACTTCTTGTATCAATTTCTTGTCTACTGAAGTGT 782
QY 904 GTGGCCAGGTGATTTCTGATCAGATTTGTAGTAAATATCATGTTACTTCTGGGTGGATTCA 963
Db 783 GCTCCGTGAGATGCTGAATCTCTTTGAGCAATGCTTATTAATTTGATGATTTG 842
QY 964 ACACCAATAATGAACATGAGCGGAGCAAGTGTATGCTGCCCACTCTTTGCTGGAA 1023
Db 843 ACAGTAAAGAACTTTAAACAGAGTGGCTTGGATGTGATCTATCACATAATCCAGCAGGAA 902
QY 1024 CATCTGTGCAAAATATTCTACCTGAGCGCAGCAGTGAATCTGTGTGAACCTCCGGCAT 1083
Db 903 CTTCGTGTTCAACATGTTCCATTTGACCCAGCCTGTATGCTGGAAATTTCAAGCTT 962
QY 1084 TTGACTGGGGAGTGAGCAACAAAATCTGGAAAATGCAATCAGCCAACTCTCTGTAAAGT 1143
Db 963 ATGACTGGGGAAGCCAGTTTCAAGATAGGATGCACTATGATCAGTCCCACTCCCTACT 1022
QY 1144 ACAGAGTCAGATATGACCGTCCCTACAGCAATGTTGGACAGAGCTCAGGCTGGCTTT 1203
Db 1023 ACAATGTGACAGCCATGAATGTAACCAATTTGAGTGTGAAACGGTGGCAAGGACCTGTTGG 1082
QY 1204 CAAATCCAGAGAGCTGAAAATGCTCTCTGAGGTGACCAACCTCATCTACCAATAAGA 1263
Db 1083 CTGACCCCAAGATGTTGGCTTTTCTTCCAAAACCTCCCAATCTTATTTTACCAAGG 1142
QY 1264 ATATTCCTGATGGGCTCAGTGAATTTCACTGGGTTTGGATGCTCTCCCTCAGCGTATGT 1323
Db 1143 AGATTCCTTTTACAATCAGTGAATTTTCACTGAGCTTTATCTGGGCAATGGAATGGCCCTCAAGAAGTTT 1202
QY 1324 ACAATGAAATCATCCATCTGAT 1345
Db 1203 ACAATGACATTTTCTATGAT 1224

RESULT 2
US-09-186-489-3
; Sequence 3, Application US/09186489
; Patent No. 6375947
; GENERAL INFORMATION:
; APPLICANT: Bolen, Paul L.
; APPLICANT: Cihak, Paul L.
; APPLICANT: Scharpf Jr., Lewis G.
; TITLE OF INVENTION: Purified Recombinant Kid Pregastric Esterase, and
; TITLE OF INVENTION: Processes for its Production and Use
; FILE REFERENCE: 5499/3

; CURRENT APPLICATION NUMBER: US/09/186,489
; CURRENT FILING DATE: 1998-11-05
; NUMBER OF SEQ ID NOS: 7
; SOFTWARE: PatentIn ver. 2.0
; SEQ ID NO 3
; LENGTH: 1411
; TYPE: DNA
; ORGANISM: Kid (Goat)
; US-09-186-489-3

Query Match 22.3%; Score 456.8; DB 4; Length 1411;
Best Local Similarity 62.0%; Pred. No. 8.4e-121;
Matches 741; Conservative 0; Mismatches 452; Indels 3; Gaps 1;
QY 167 AATGGAATCTGGCTTCTGATTTCTGGTGGCTATATGTTCCAGAGAAATGTGAATTCAGT 226
Db 61 ATTGAATATGTGGTGGCTACTTGTAAAGGTGTGTTTCATCCACATGCTGGAATGCATT 120
QY 227 ACATATGC---CAACTAAGCTGTGACCCAGAGCAATTCATGAATATTAGTGAATCAT 283
Db 121 TTGTTTCTTGGAAAATTTCTAAGAACCTTGAAGCCAGTATGAATGTGAGTCAGATGAT 180
QY 284 CCAACATCAAGGCTATCCCTGTGAGGAATATGAAGTGGCAACTGAAGATGGGTATATCCT 343
Db 181 TTCTTCTGGGCTTACCAAGTGAAGTGAATAAGTTATTAATCTGAGATGGCTATATCCT 240
QY 344 TTCTGTTAAGAGGATTCCTCGAGGCTTAGTGCAACCTAAGAGACAGGTTCCAGGCTGT 403
Db 241 TCAGGTCTATCGGATTCCTCATGAAAGAAATGATGCTTAATCATTTAGGTGAGAGACTGT 300
QY 404 GGTGTTACTGAGCATGGCTTAGTTGGAGTGTGCAACTGAGTATTCACAACTGCCAA 463
Db 301 TGTGTTTCTGAGCAATGGTCTTCTTGCTCAGCTACAACTGAGTTTCCAACTTCCAA 360
QY 464 CAATAGCTGGCTTCAATTTCTGAGATGCTGTTTGTGAGCTGTGAGTGGGAAACAGCAG 523
Db 361 CAACAGCTGGGCTTCTCTCTGAGCATGCTGTTTATGACGTGTGGCTGGGAAACAGCAG 420
QY 524 GGGAAACGCTGCTCGAAAACACAGACACTCTCATAGACCAAGATGAGTTCTGGGC 583
Db 421 AGGAAACACTTGGGCCAGGAACTTTATTAATTAATTAATTAATTAATTAATTAATTAAT 480
QY 584 TTTCAAGTTATGATGAGATGGCTAGTTTGAACCTTCTGAGTGAATAAACTTTATTTGCA 643
Db 481 TTTCAAGTTTATGATGAGATGGCTAGTTTGAACCTTCTGAGTGAATAAACTTTATTTAA 540
QY 644 GAAACCGGCCAGGAAAGATCTATTATGCTGGCTATTGAGGCTTACAGGCGACCACTGGGCT 703
Db 541 GAGAACAGGACAGAGAGCTACACTATGTTGGCCATTCACAGGCGACCACTGGGCTTT 600
QY 704 TATTGCAATTTTCCACCATGCGCAGAGCTGGCTCAGAAAATCAAAATGATTTTGTCTTTC 763
Db 601 TGTGGCTTTTCTACCAATCCCACTGGCTGAAAATTAATTAATTAATTAATTAATTAAT 660
QY 764 ACCCATAGCACTGTTAAGCATGCAAAAGCCCGGACCAAAATTTTGTGTTGTCGCA 823
Db 661 CCCAGTCCCACTGAGTGAAGCACCACCGAGGCTGTTTAAACAACTTGCATTTATTTCTCA 720
QY 824 TATGATGATCAAGGATTTGTTGGCAAAAAGAAATTTCTGTATCAGACCAAGATTTCTCAG 883
Db 721 CTTCTCTCTTCAAGATTATATTGTTAACAAATGTTCTACCCACACAAATTTTTTGAACA 780
QY 884 ACACTTCTGTTATTAATCTTTTGGCCAGGTGATTTCTTGTATCAGATTTGTTAGTATCAT 943
Db 781 ATTTCTTGGTGTGAAGTGTCTCTCTGAGACACTGAGTGTCTCTTTTGTGAAGATGCTTT 840
QY 944 GTTACTTCTGGTGGATTCACACCAATATGAACATGAGCGGAGCAAGTGTATATGC 1003
Db 841 GTTTCCTTACTGAGGCTGACATTAATAAATTTCAACATGAGTGTCTGATGATGTATGT 900
QY 1004 TGCCCACTCTTGTGGAACATCTGTGCAAAATTTCTACACTGGAGCGGAGGAGTGA 1063
Db 901 AGCACATAATCCAGCAGGAGCTTCTGTTTCAAAAACATCTCTCCACTGGAGAGAGGCTATTA 960

QY 1064 TTCTGGTGAATCCGGGCAATTTGACTGGGAGTGAGACCAAAATCTGGAATAATGCAA 1123
Db 961 GTCTGGGAATTCAGCTTTTGACTGGGAGCCTCAGTTGAGAACCTTAATGCAATTATA 1020
QY 1124 TCAGCCAACTCCTGTGAAGGTACAGAGTACAGAGATATGACGGTCCCTACAGCAATGTGGAC 1183
Db 1021 TCAGCCCAACCTCCCATCTACAATTTAAACAGCCATGAATGTCCCAATTCAGATGGAG 1080
QY 1184 AGGAGTTCAGACTGGCTTTCAATCCAGAGAGAGTGAATGCTCTCTGAGGTGAC 1243
Db 1081 TGCTGGCCCAAGACCTGTGGGTGACCTTCAGGATGTTGACCTTTTGCTTTCAAACTCTC 1140
QY 1244 CAACCTCATCTACCAATGAATAATTCCTGAATGGGCTCACGTGGATTTCACTCTGGGGTTT 1303
Db 1141 TAATCTCATTCACCAAGGAATTCCAAATTAACATCATCTGGAATTTATCTGGGCAAT 1200
QY 1304 GGATGCTCTCACCGTATGTAATGAATCATCCATCTGATGAGCAGAGGAGA 1359
Db 1201 GGATGCACTCAAGAGTTTACAATGAATTTATTTCTTGATGGCAAAAGACAAA 1256

RESULT 3

US-10-043-665B-3
; Sequence 3, Application US/10043665B
; Patent No. 6582948
; GENERAL INFORMATION:
; APPLICANT: Bolen, Paul L.
; APPLICANT: Cihak, Paul L.
; APPLICANT: Scharpf Jr., Lewis G.
; TITLE OF INVENTION: Recombinant Kid Pregastric Esterase and Methods for Its
; TITLE OF INVENTION: Production and Use
; FILE REFERENCE: IFF-0009
; CURRENT APPLICATION NUMBER: US/10/043.665B
; CURRENT FILING DATE: 2002-09-24
; PRIOR APPLICATION NUMBER: US 09/186,489
; PRIOR FILING DATE: 1998-11-05
; NUMBER OF SEQ ID NOS: 7
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 3
; LENGTH: 1411
; TYPE: DNA
; ORGANISM: Kid (Goat)
US-10-043-665B-3

Query Match 22.3%; Score 456.8; DB 4; Length 1411;
Best Local Similarity 62.0%; Pred. No. 8.4e-121;
Matches 741; Conservative 0; Mismatches 452; Indels 3; Gaps 1;
QY 167 AATGAAATGTGGCTTCTGATTTCTGGTGGCGTATATGTTCCAGAGAAATGTGAATTCAGT 226
Db 61 ATTGAATGTGGTGGCTTCTGATTTGTAACGGTGTGTTTCAATCCATGTCGGAATGCATT 120
QY 227 ACATATGC---CAACTAAAGCTGTGGACCCAGAGAGCATTCATGAATATTAATGAATCAT 283
Db 121 TTGTTCTCTGGAAAAATTTGTAAGAACCCCTGAAGCCAGCATGTAATGTGAGTCAGATGAT 180
QY 284 CCAACATCAAGCTATCCCTGTGAGGAATATGAAGTGGCACTCAAGATGGGTATATTCCT 343
Db 181 TTCTTCTGGGGCTACCAAGTGAAGTGAATGATTAATGATGATGATGATGATGATGATGAT 240
QY 344 TTCTGTTAAAGATTTCTCGAGGCTTAGTGCAACCTTAAGAAAGACAGGTTCCAGGCTGT 403
Db 241 TCAGGTCTATCGGATTTCTCATGGAAGAAATGATGCTTAATCATTTAGGTTCAGAGACCTGT 300
QY 404 GGTGTTACTGAGCATGGCTTAGTTGGAGGTGCTAGCACTGGATTTCCAACTGGCCCAA 463
Db 301 TGTGTTTCGAGCATGCTGCTTCTGCTGAGCTACAACTGGAATTTCCAACTTTCCCAA 360
QY 464 CAATAGCTGGCTTCATTTCTGGCAGATGCTGTTTTCAGGTGTGATGGGGAACAGCAG 523
Db 361 CAACAGCTGGCTTCTCTCTGGCAGATGCTGTTTATGACGTGTGGGTGGGAACAGCAG 420

QY 524 GGGAAACGCTGTCTCGAAACACACAGACACTCTCCATAGACCAAGATGAGTTCTGGGC 583
Db 421 AGGAAACACTTTGGGCCCCAGGAAACATTTATATTAATTAACAGACTCCCTGAAATCTGGGC 480
QY 584 TTTTCAGTTATGATGAGATGGCTTAGGTTGACCTTCTGCAAGTGAATAAACTTTATTTTGA 643
Db 481 TTTTCAGCTTTGATGAATGGCTTGAATATGACCTTCCATCTACAAATGATTTTCATCTTAA 540
QY 644 GAAACGGCCACAGGAAGATCTATATGTGGCTTATTCAGGGCAACCATGGGTT 703
Db 541 GAAACAGACAGAGAAAGCTACATATGTTGGCCATTTCCCAAGGCAACCATTTGGTT 600
QY 704 TATTGCAATTTCCACATGCCAGAGCTGCTCAGAAAAATCAAAATGATTTTGTCTTTAGC 763
Db 601 TGTGSCCTTTTCTACCAATCCACACTGGCTGGTGAAGAAATCGAGTCTTCCATGCAATAGC 660
QY 764 ACCATAGCCACTGTTAAGCATGAAAGCCCCGGGACCAAAATTTTGTGTGCTCCAGA 823
Db 661 CCCAGTGGCCACAGTGAAGCACACCCAGAGCCTGTTTAAACAAACTTGCATTTATTCCTCA 720
QY 824 TATGATGATCAAGGCAATTTTGGCAAAAGAAATTTCTGTATCAGACCAAGTTTCTCAG 883
Db 721 CTTCCTCTTCAGATATATTTGTTAAACAAATGTTCTACCCACAAATTTTGAACA 780
QY 884 ACAACTTGTATTATACCTTTTGGCCAGGTGATTTCTGATCAGATTTGTAGTAATATCAT 943
Db 781 ATTTCTTGGTGTGAAGTGTCTCTCGTGAGACACTGGATGTCTTTTAAAGATGCTT 840
QY 944 GTTACTTCTGGGTGATTTCAACACCAACAAATATGAACATGAGCCGAGCAAGTGTATATGC 1003
Db 841 GTTGGCCATTACTGAGCTGACATPAAAACTTCAACATGAGTCTGCTTGAATGTATGT 900
QY 1004 TGCCCACTCTTGTGTGAACATCTGTCAAAATATTTCTACCTGAGCCAGGCAAGTCAA 1063
Db 901 AGCATAATATCCAGCAGGAGCTTTGTTCAAAACATCTCTCAGTGAGACAGGCTATTAA 960
QY 1064 TTCTGTGAACTCCGGGCAATTTGACTGGGGAGTGAGACCAAAATCTGGAATAATGCAA 1123
Db 961 GTCTGGAAATTCAGGCTTTGACTGGGGAGCTCAGTGAACCTTAATGCAATATAA 1020
QY 1124 TCAGCCAACTCCTGTGAAGTACAGAGTACAGATATGACGGTCCCTACAGCAATGTGAC 1183
Db 1021 TCAGCCCAACCTCCATCTACAAATTTAAACAGCCATGAATGTCCCAATTTGCAATGAG 1080
QY 1184 AGAGGTGAGACTGGCTTTCAAAATCCAGAGAGCTGAAATGCTGCTCTGAGGTGAC 1243
Db 1081 TGCTGGCCAAAGACCTGTGGCTGACCTCAGGATGTTGACCTTTGCTTTCAAACTCTC 1140
QY 1244 CAACCTCATCTACCAATAAGAAATTTCTGAAATGGGCTCACGTGGATTTCACTGGGGTTT 1303
Db 1141 TAATCTCATTCACCAAGGAATTTCCAAATTAACATCATCTGGAATTTATCTGGGCAAT 1200
QY 1304 GGATGCTCTACCTATGTAATGTAATGAATCATTCATCTGATGAGCAGGAGGAGA 1359
Db 1201 GGATGCACTCAAGAGTTTACAATGAATTTATTTCTTGATGGCAAAAGACAAA 1256

RESULT 4

US-09-186-489-4
; Sequence 4, Application US/09186489
; Patent No. 6375947
; GENERAL INFORMATION:
; APPLICANT: Bolen, Paul L.
; APPLICANT: Cihak, Paul L.
; APPLICANT: Scharpf Jr., Lewis G.
; TITLE OF INVENTION: Purified Recombinant Kid Pregastric Esterase, and
; TITLE OF INVENTION: Processes for Its Production and Use
; FILE REFERENCE: 5499/3
; CURRENT APPLICATION NUMBER: US/09/186,489
; CURRENT FILING DATE: 1998-11-05
; NUMBER OF SEQ ID NOS: 7
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 4

```

; LENGTH: 1134
; TYPE: DNA
; ORGANISM: Bovine
US-09-186-489-4

Query Match      22.2%; Score 453.6; DB 4; Length 1134;
Best Local Similarity 62.7%; Pred. No. 6.1e-120;
Matches 705; Conservative 0; Mismatches 419; Indels 0; Gaps 0;

QY 236 AACTAAAGCTGTGGAGCCAGAGCAATTCATGAATATTAGTGAATCATCAATCAATCAAGG 295
Db ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
QY 9 AAAAAATGTGAAGACCCCTGAAGCCAGTATGAATGTTAGTCAGATGATTTCTACTGGGG 68
Db ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||

QY 296 CTATCCCTGTGAGGAATATCAAGTCGCAACTGAAGATGGTATATCTCTTCTGTTAAACAG 355
Db ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||

QY 69 CTACCAAGTGAGATGCATTAAGTATTAATCTGGGATGGTATATCTCTCAGGTCTATCG 128
Db ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||

QY 356 GATTCCTCGAGGCTGTAGTGAACCTTAAGAGACAGGTTCCAGGCTGTGGTGTACTGCA 415
Db ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||

QY 129 GATTCCTCATGGAAGAATAATCTAATCTATTTAGGTCAGAGACCTGTGTGTTCTGCA 188
Db ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||

QY 416 GCATGGCTAGTTGGAGGCTGTAGCAACTGAGTATTCGAATCTCCCAACATAGACCTGGG 248
Db ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||

QY 189 GCATGGCTCTCTTGGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 248
Db ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||

QY 476 CTTCATTTCTGCGAGATCTGTTTTCAGCTGTGGTGGGAAACAGCAGGGGAAACGCTG 535
Db ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||

QY 249 CTTCCTCTGGCAGATGCTGGTATGATGATGATGATGATGATGATGATGATGATGATG 308
Db ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||

QY 536 GTCTCGAAAAACACAAGACACTCTCCATAGAACAGATGATGATGATGATGATGATGATG 595
Db ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||

QY 309 GGCCCGAGGAACATTTATCTATTTACCAAGACTCCCGGAAATCTGGGCTTTTCACTTGA 368
Db ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||

QY 596 TGACATGCTAGGTTTTCAGCTCTCTCTGAGTGAATAAATTTATTTTTCAGAAAAACGGCCA 655
Db ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||

QY 369 TGAATGGGGAATATGACCTTCCATCTACATTTGATTTTCAATTAGAGAGAACAGACA 428
Db ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||

QY 656 GGAAGAAGATCTATTATGTCGGCTATTTCAGAGGACCAACCATGGCTTTTATTCGATTTTC 715
Db ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||

QY 429 GAAGAAGCTACATATGTTGGCCATTTCCCAAGGACCAACATTTGTTTATCGCCTTTTC 488
Db ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||

QY 716 CACCATGCCAGAGCTGCTCAGAAAAATCAAAATGATTTTGTCTTTAGCACCATAGCCAC 775
Db ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||

Query Match      22.2%; Score 453.6; DB 4; Length 1134;
Best Local Similarity 62.7%; Pred. No. 6.1e-120;
Matches 705; Conservative 0; Mismatches 419; Indels 0; Gaps 0;

QY 236 AACTAAAGCTGTGGAGCCAGAGCAATTCATGAATATTAGTGAATCATCAATCAATCAAGG 295
Db ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||

QY 9 AAAAAATGTGAAGACCCCTGAAGCCAGTATGAATGTTAGTCAGATGATTTCTACTGGGG 68
Db ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||

QY 296 CTATCCCTGTGAGGAATATCAAGTCGCAACTGAAGATGGTATATCTCTTCTGTTAAACAG 355
Db ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||

QY 69 CTACCAAGTGAGATGCATTAAGTATTAATCTGGGATGGTATATCTCTCAGGTCTATCG 128
Db ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||

QY 356 GATTCCTCGAGGCTGTAGTGAACCTTAAGAGACAGGTTCCAGGCTGTGGTGTACTGCA 415
Db ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||

QY 129 GATTCCTCATGGAAGAATAATCTAATCTATTTAGGTCAGAGACCTGTGTGTTCTGCA 188
Db ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||

QY 416 GCATGGCTAGTTGGAGGCTGTAGCAACTGAGTATTCGAATCTCCCAACATAGACCTGGG 248
Db ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||

QY 189 GCATGGCTCTCTTGGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 248
Db ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||

QY 476 CTTCATTTCTGCGAGATCTGTTTTCAGCTGTGGTGGGAAACAGCAGGGGAAACGCTG 535
Db ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||

QY 249 CTTCCTCTGGCAGATGCTGGTATGATGATGATGATGATGATGATGATGATGATGATG 308
Db ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||

QY 536 GTCTCGAAAAACACAAGACACTCTCCATAGAACAGATGATGATGATGATGATGATGATG 595
Db ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||

QY 309 GGCCCGAGGAACATTTATCTATTTACCAAGACTCCCGGAAATCTGGGCTTTTCACTTGA 368
Db ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||

QY 596 TGACATGCTAGGTTTTCAGCTCTCTCTGAGTGAATAAATTTATTTTTCAGAAAAACGGCCA 655
Db ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||

QY 369 TGAATGGGGAATATGACCTTCCATCTACATTTGATTTTCAATTAGAGAGAACAGACA 428
Db ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||

QY 656 GGAAGAAGATCTATTATGTCGGCTATTTCAGAGGACCAACCATGGCTTTTATTCGATTTTC 715
Db ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||

QY 429 GAAGAAGCTACATATGTTGGCCATTTCCCAAGGACCAACATTTGTTTATCGCCTTTTC 488
Db ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||

QY 716 CACCATGCCAGAGCTGCTCAGAAAAATCAAAATGATTTTGTCTTTAGCACCATAGCCAC 775
Db ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||

```

Db 489 TACCAGTCCACATGGCTGCTGAAATCAAGTCTCTATGATAGCCCGAGTGGCCAC 548
 QY 776 TGTTAGCATGCAAAAGCCCGGACCAAAATTTTGTGCTGCGCAGATATGATGATCA 835
 Db 549 AGTGAAGTACACAGAGCGCTGTTAAACAACATGCACTATTCTCTCACTTCTTCAA 608
 QY 836 GGGATTGTTGGCAAAAGAAATTTCTGTATCAGACGAGATTTCTCAGACAACTTGTAT 895
 Db 609 GATTATATTGGTGACAAATGTTCTACCCACACATTTTGGAAACATTTCTTGGTGT 668
 QY 896 TTACCTTTGGCCAGGTGATTTCTGATCAGATTTCTAGTAATATCATCTTACTTCTGGG 955
 Db 669 TGAATATGTCCTCCGTGAGACATGCGATGTCCTTTCTGAAGAATGCCCTTTGGCCATTAC 728
 QY 956 TGGATTCAACACCAACAAATATGAACATGAGCCGAGCAAGTGTATATCTGCGCCACACTCT 1015
 Db 729 TGGATGTGCAATTAATACTTCAACATGATGCTCTAGATGTTGTATATAGCAATATCC 788
 QY 1016 TGTGGAACATCTGTGCAAAATATTTCTACATGAGCCAGGAGTGAATTTCTGGTGAAT 1075
 Db 789 AGCAGGAATCTTGTCTCAAAACACCTCCACTGGACAGAGGCTGTTAAGTCTGGAAAT 848
 QY 1076 CCGGCATTTGACTGGGGAGTGCAGACCAAAATCTGGAATAATGCAATCAGCCACTCC 1135
 Db 849 CCAAGCTTTGACTGGGGAGCCCATATCAGAACCTTAATGCAATATCATCAGCCACACC 908
 QY 1136 TGTAAGGTACAGAGTACAGATATGACGGTCCCTACAGCAATGTGGACAGAGGTGAGGA 1195
 Db 909 TCCCATCTCAATTAACAGCCATGAATGTCCCAATTCAGATGAGGTGCTGACATGA 968
 QY 1196 CTGGCTTTCAATCCAGAGACGTGAATTTCTATGAGTGGGTTTGGATGCTCTCA 1255
 Db 969 CTGTTGGCTGACCCCTCAGGATGTTGACCTTTTGTCTTCAAAAACCTCTCTAATCTCAT 1028
 QY 1256 CCATAAGAATATTTCTGAAATGGGCTCAGCTGGATTTTCTATCTGGGGTTTGGATCTCTCA 1315
 Db 1029 CCACAAGGAATTTCCAAATTAACAATCATCTGGACTTTATCTGGGCAATGATGACCTCA 1088
 QY 1316 CCGTATGTACATGAATCATCCATCTGATGACGAGGAGGAGA 1359
 Db 1089 AGAAGTTTCAATGAATGAATTTCTTCTTGTGAGCCGAGCAAAA 1132

RESULT 6

US-09-186-489-1
 ; Sequence 1, Application US/09186489
 ; Patent No. 6375947
 ; GENERAL INFORMATION:
 ; APPLICANT: Bolen, Paul L.
 ; APPLICANT: Cihak, Paul L.
 ; APPLICANT: Scharpf Jr., Lewis G.
 ; TITLE OF INVENTION: Purified Recombinant Kid Pregastric Esterase, and
 ; FILE OF INVENTION: Processes for its Production and Use
 ; FILE REFERENCE: 5499/3
 ; CURRENT APPLICATION NUMBER: US/09/186,489
 ; CURRENT FILING DATE: 1998-11-05
 ; NUMBER OF SEQ ID NOS: 7
 ; SOFTWARE: Patent in Ver. 2.0
 ; SEQ ID NO 1
 ; LENGTH: 1134
 ; TYPE: DNA
 ; ORGANISM: Kid (Goat)
 ; FEATURE:
 ; NAME/KEY: CDS
 ; LOCATION: (1) . . . (1134)
 ; US-09-186-489-1

Query Match 22.0%; Score 450.4; DB 4; Length 1134;
 Best Local Similarity 62.5%; Pred. No. 5.1e-119;
 Matches 703; Conservative 0; Mismatches 421; Indels 0; Gaps 0;
 QY 236 AACTAAGCTGTGGACCCAGAGCAATCATGATATTAATGTAATCATCCAAATCAAGG 295

Db 9 AAAAATTTGCTAAGAACCTTGAAGCCAGTATGAATGTGAGTCAGATGATTTCTTCTGGGG 68
 QY 296 CTATCCCTGTGAGGAATATGAAGTGCACACTGAAGATGGGTATATCTTTCTGTAAACAG 355
 Db 69 CTACCCAAAGTGAATGATGAATGAATTAACCTGAGATGGCTATATCTTCAAGGTCTATCG 128
 QY 356 GATTCTCTGAGGCTTAGTGCAACCTAAGAAGACAGGTTCCAGGCTGTGGTGTACTGTCA 415
 Db 129 GATTCTCTCATGGAAGAATGATGCTAATCTAGTTCAGAGACCTGTTGTGTCTTGTGCA 188
 QY 416 GCATGGCTTAGTTGGAGGTGCTAGCAACTGGATTTCCAACTGCCCCAACAAATAGCCCTGGG 475
 Db 189 GCATGGCTCTTCTTGGCTCAGCTACAACTGGAATTTCCAACTTCCCAACAACAGCCCTGGG 248
 QY 476 CTTCATTTCTGGCAGATCTGGTTTGTGAGCTGTGGATGGGAAACAGAGGGGAAACGCTG 535
 Db 249 CTTCCTCTGCGCAGATCTGGTTATGACGTGTGGTGGGAAACAGAGAGAGAAACACTTG 308
 QY 536 GTCTCGAAAACACAAAGACACTCTCCATAGACCAAGATGAGTCTTGGGCTTTCAAGTTATGA 595
 Db 309 GGGCCAGGACATTTTACTATTACCCAGACTCCCTGGAATTTCTGGCTTTTCAAGCTTTGA 368
 QY 596 TGAGATGGCTAGTTGACCTTCTGCTGAGTGATAAATTTTATTTTGCAGAAAAACGGGCCA 655
 Db 369 TGAATGGCTGAATATGACCTTCCATCTACAATTTGATTTTCACTTAAAGAGAAACAGGACA 428
 QY 656 GGAAGAAGATCTATTATCTCGGCTATTACAGAGGACCAACCATGCGCTTTTATTCGATTTTC 715
 Db 429 GAAGAAGCTACATATGTTGGCCATTTCCAAAGGACCAACCATTTGGTTTGTGCGCTTTTC 488
 QY 716 CACCATGCCAGAGCTGGCTCAGAAAATCAAAATGTATTTTGTCTTTAGCACCCATAGCCAC 775
 Db 489 TACCAATCCACACATGCTGCTGAAAATTCGAAGTCTTCCATGATGATAGCCAGTGGCCAC 548
 QY 776 TGTTAAGCATGCAAAAGACCCCGGACCAATTTTGTGCTGCGCAGATATGATGATCA 835
 Db 549 AGTGAAGACACCCAGAGCTGTTTAAACAATTTGCACTTATTTCTCTCACTTCTTCAA 608
 QY 836 GGGATTTCTTGGCAAAAGAAATTTCTGTATCAGACGAGATTTCTCAGACAACTTGTAT 895
 Db 609 GATTATATTGGTAAACAAATGTTCTACCCACACAAATTTTGTGAACAAATTTCTTGTGTGT 668
 QY 896 TTACCTTTGGCCAGGTGATTTCTGATCAGATTTGTAGTAATATCATGTTACTTCTGGG 955
 Db 669 TGAAGTGTGCTCTGCTGAGACACTGGATGTCCTTTTGAAGAATGCTTGTGTGCCATTAC 728
 QY 956 TGGATTCAACACCAACAAATATGAACATGAGCCGAGCAAGTGTATATGCTGCCACACTCT 1015
 Db 729 TGGAGCTGACATTAATAAATTTCAACATGAGTCTGTTAGTGTATGTAGCACATATCC 788
 QY 1016 TGCTGGAACATCTGTGCAAAATATTTCTA CACTGAGCCAGGACAGTGAATTTCTGGTGAAT 1075
 Db 789 AGCAGGAGCTTCTGTTCAAAACATCTCTCCACTGAGACAGGCTATTAAAGTCTGGGAAAT 848
 QY 1076 CCGGCATTTGACTGGGGAGTGCAGACCAAAATCTGGAATAATGCAATCAGCCACTCC 1135
 Db 849 CCAAGCTTTGACTGGGGAGCCCTCAGTTGAGAACCTTAATGCAATATATCAGCCCAACC 908
 QY 1136 TGTAAGGTACAGAGTACAGATATGAGGTCCTTACAGCAATGTGGACAGGAGGTGAGGA 1195
 Db 909 TCCCATCTCAAAATTTAAGCCATGAATGTCCCAATTTGCAAGTATGAGTGTGCGCAAGA 968
 QY 1196 CTGGCTTTCAAAATCCAGAGACGTGAATGCTGCTCTGAGGTGACCAACTCATCTA 1255
 Db 969 CTGTTGGCTGACCCCTCAGGATGTTGACCTTTTGTCTTCAAAAACCTCTCTAATCTCAT 1028
 QY 1256 CCATAAGAATATTTCTGAAATGGGCTCAGCTGGATTTTCTATCTGGGGTTTGGATCTCTCA 1315
 Db 1029 CCACAAGGAATTTCCAAATTAACAATCATCTGGACTTTATCTGGGCAATGATGACCTCA 1088
 QY 1316 CCGTATGTACATGAATCATCCATCTGATGACGAGGAGGAGA 1359

Db 1089 AGAAGTTTACATGAATTAATTTCTTGTATGCGCAAGACAA 1132

US-10-043-665B-1
 ; Sequence 1, Application US/10043665B
 ; Patent No. 6582948
 ; GENERAL INFORMATION:
 ; APPLICANT: Bolen, Paul L.
 ; APPLICANT: Cihak, Paul, L.
 ; APPLICANT: Scharpf Jr., Lewis G.
 ; TITLE OF INVENTION: Recombinant Kid Pregastric Esterase and Methods for Its
 ; TITLE OF INVENTION: Production and Use
 ; FILE REFERENCE: IFF-0009
 ; CURRENT APPLICATION NUMBER: US/10/043,665B
 ; CURRENT FILING DATE: 2002-09-24
 ; PRIOR APPLICATION NUMBER: US 09/186,489
 ; PRIOR FILING DATE: 1998-11-05
 ; NUMBER OF SEQ ID NOS: 7
 ; SOFTWARE: PatentIn Ver. 2.0
 ; SEQ ID NO 1
 ; LENGTH: 1134
 ; TYPE: DNA
 ; ORGANISM: Kid (Goat)
 ; FEATURE:
 ; NAME/KEY: CDS
 ; LOCATION: (1)..(1134)
 US-10-043-665B-1

Query Match 22.0%; Score 450.4; DB 4; Length 1134;
 Best Local Similarity 62.5%; Pred. No. 5.1e-119;
 Matches 703; Conservative 0; Mismatches 421; Indels 0; Gaps 0;

Qy 236 AACTAAAGCTGTGGACCCAGCAATTCATGAATATTAGTGAATCATCAACATCAAGG 295
 Db 9 AAAAAATGCTAAGAACCTCGAAGCCAGTATGAATGTGAGTCAGATGATTTCTTGGGG 68

Qy 296 CTATCCCTGTGAGGAATATGAGTCGCAACTGAGATGGGTATATCTTCTGTTAAACAG 355
 Db 69 CTACCCAGTGAATGATGAATGAATGAATGAATGAATGAATGAATGAATGAATGAAT 128

Qy 356 GATTCTCTGAGGCTAGTGCACCTAAGAAAGACAGAGTTCCAGGCTGTGGTGTACTGCA 415
 Db 129 GATTCTCTGAGGCTAGTGCACCTAAGAAAGACAGAGTTCCAGGCTGTGGTGTACTGCA 188

Qy 416 GCATGGCTAGTGGAGGTGCTGCAACTGATTTCCAACTGCGCCCAACATAGCTGGG 475
 Db 189 GCATGGCTAGTGGAGGTGCTGCAACTGATTTCCAACTGCGCCCAACATAGCTGGG 248

Qy 476 CTTCAATCTGCGAGATGCTGTTTGTGAGTGTGAGTGGGAAACAGAGGGAACGCTG 535
 Db 249 CTTCTCTGCGAGATGCTGTTTGTGAGTGTGAGTGGGAAACAGAGGGAACGCTG 308

Qy 536 GTCTCGAAACACAAAGACATCTCTCCATAGACCAAGATGATTTCTGGGCTTTCAGTTATGA 595
 Db 309 GGCCCGAGGAACATTTAT 368

Qy 596 TGAGATGCTAGTGTGACCTTCTGAGTGAATAAATTTATTTTGCAGAAACAGGCGCA 655
 Db 369 TGAATGCTGAATGATGACCTTCTGAGTGAATAAATTTATTTTGCAGAAACAGGCGCA 428

Qy 656 GGAAGAAGCTATTATGTCGCTATTCAAGGCAACCAACATGCGGCTTTATTCATTTTC 715
 Db 429 GAAGAAGCTATTATGTCGCTATTCAAGGCAACCAACATGCGGCTTTATTCATTTTC 488

Qy 716 CACCATGCGAGGCTGCTGAGAAATCAGAAATGATTTTGTGCTTACGACCAATAGCCAC 775
 Db 489 TACCATGCGAGGCTGCTGAGAAATCAGAAATGATTTTGTGCTTACGACCAATAGCCAC 548

Qy 776 TGTAAAGCATCAAAAAGCCCGGACCAAAATTTTGTGCTGCGAGATGATGATCA 835
 Db 549 AGTGAAGCAGACCCAGAGGCTTTTAAACAACTTGCACTTATCTCTCACTTCTCTTCAA 608

Qy 836 GGGATTGTTGGCAAAAAGAAATTTCTGTATCAGACAGATTTCTCAGACAACTTGTAT 895
 Db 609 GATTATATTTGGTAAACAAAATTTCTATCCCAACAAATTTTGTGAAATTTCTTGTGT 668

Qy 896 TTACCTTTGTCGCGAGGATTTCTGTATCAGATTTGTAGTAATATCATGTTACTTCTGGG 955
 Db 669 TGAAGTGTCTCTGTCGAGACATGATGCTTTTGAAGAATGCTTTGTTGCCATTAC 728

Qy 956 TGGATTCAACACCAACAAATATGAACATGAGCGGAGCAAGTGTATATGCTGCCCACTCT 1015
 Db 729 TGGAGCTGACAAATAAAATTTCAACATGAGTGTGATGATGATGATGATGATGATGAT 788

Qy 1016 TGCTGGAACATCTGTGCAAAATATTCTACACTGAGCGAGCACTGAAATCTGTGGAAT 1075
 Db 789 AGCAGGAGCTTCTGTTCAAAACATCTCCACTGAGACAGGCTATTAAAGTCGGAAAT 848

Qy 1076 CCGGCAATTTGACTGGGGAGTGAGACCAAAATCTGGAATAAATGCAATCAGCAACTCC 1135
 Db 849 CCAAGCTTTTGAATGGGAGCTCAGTTGAGAACCTAATGCAATTAATCAGCCCCACCC 908

Qy 1136 TGTAAGGTACAGAGTCAGAGATATGACGTCCTACAGCAATGTGACAGAGGTCAAG 1195
 Db 909 TCCCATCTCAAAATTTACAGCCATGATGTCCTCAATTTGAGTATGAGTGTGAGTGTG 968

Qy 1196 CTGGCTTTCAAAATCCAGAAAGCTGAAATGCTCTCTGAGGTGACCAACTCATCTA 1255
 Db 969 CTGTGCTGACCTCAGGATGTTGACCTTTTCTTCAAACTCTCTAATCTCAATCA 1028

Qy 1256 CCATGAAGATATCTCTGAATGGGCTCAGTGGATTTCTATCTGGGTTTGGATGCTCTCA 1315
 Db 1029 CCACAGGAATTTCAAAATTAACAATCTGGAATTTATCTGGCAATGATGATGATGAT 1088

Qy 1316 CCGTATGTACAAATGAATCATCCATCTGATGAGCGAGGAGGA 1359
 Db 1089 AGAAGTTTACAAATGAATTAATTTCTTGTATGCGCAAGACAA 1132

RESULT 8
 US-09-186-489-5
 ; Sequence 5, Application US/09186489
 ; Patent No. 6375947
 ; GENERAL INFORMATION:
 ; APPLICANT: Bolen, Paul L.
 ; APPLICANT: Cihak, Paul L.
 ; APPLICANT: Scharpf Jr., Lewis G.
 ; TITLE OF INVENTION: Purified Recombinant Kid Pregastric Esterase, and
 ; FILE REFERENCE: 5499/3
 ; CURRENT APPLICATION NUMBER: US/09/186,489
 ; CURRENT FILING DATE: 1998-11-05
 ; NUMBER OF SEQ ID NOS: 7
 ; SOFTWARE: PatentIn Ver. 2.0
 ; SEQ ID NO 5
 ; LENGTH: 8324
 ; TYPE: DNA
 ; ORGANISM: Yeast YE-1 expression vector
 US-09-186-489-5

Query Match 22.0%; Score 449.8; DB 4; Length 8324;
 Best Local Similarity 62.7%; Pred. No. 2.2e-118;
 Matches 700; Conservative 0; Mismatches 417; Indels 0; Gaps 0;

Qy 243 GCTGTGGACCCAGAGCAATTCATGAATATTAGTGAATATCATCAACATCAAGGCTATCCC 302
 Db 1547 GCTAAGAACCTTGAAGCCAGTATGAATGTGAGTCAGATGATTTCTTCTGGGCTACCCA 1606

Qy 303 TGTGAGGAATATGAAGTGCAGCACTGAGATGGGTATATCTTCTGTAAACAGATTCCT 362
 Db 1607 AGTGAGTGCATAAAGTTATAACTGCAGATGGCTATATCTTCTGAGGTCTATCGGATTCCT 1666

Qy 363 CGAGGCTAGTGCACCTTCAAGAACAGGTTCCAGGCTGTGGTGTACTGACAGATGCG 422
 Db 1667 CATGGAAGATGATGCTAATCAATTTAGTTCAGAGACCTGTTGTGTTCTGACAGATGGT 1726

423 CTAGTTGAGGTGCTAGCACTGATTTCCAACTGCTCCCAACATACCTGCTGGCTTCATT 482
 1727 CTTCTTGCTCAGTACCAACTGATTTCCAACTGCTCCCAACATACCTGCTGGCTTCCTC 1786
 483 CTGGCAGATGCTGTTTGTGATGCTGATGGGAAACAGCAGGGAAGACGCTGCTGCGA 542
 1787 CTGGCAGATGCTGTTTGTGATGCTGATGGGAAACAGCAGGGAAGACGCTGCTGCGA 1846
 543 AAACACAGACACTCTCCATAGACCAAGATGATGCTGCTGCTGCTGCTGCTGCTGCTG 602
 1847 GAACTTTTACTATTACAGATCTCCCTGATCTGCTGCTGCTGCTGCTGCTGCTGCTG 1906
 603 GCTAGTTGACCTTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 662
 1907 GCTGATATGACCTTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 1966
 663 ATCTATTATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 722
 1967 CTACACTATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 2026
 723 CCAGAGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 782
 2027 CCCACACTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 2086
 783 CATGCAAAAGCCCGGAGCAAAATTTTGTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 842
 2087 CACACCCAGAGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 2146
 843 TTGGCAAAAGAAATTTCTGATCAGACCAAGATTTCTGACACAACTTGTATTACCTT 902
 2147 TTGGTAAACAAATTTCTACCCACACAAATTTTGTGCTGCTGCTGCTGCTGCTGCTG 2206
 903 TGTGGCAGGTGATTTCTGATCAGATTTGTAGTAAATATCATGTTTACTTCTGGGTG 962
 2207 TGCTCTGCTGAGACACTGATGCTCTTTGTAAGATGCTTGTGCTGCTGCTGCTGCTG 2266
 963 AACACACAAATATGAAATGAGCGGAGCAAGTGTATGCTGCCCACTCTTCTGCTGGA 1022
 2267 GACAAATAAAACCTTCAACATGAGTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 2326
 1023 ACATCTGTGCAAAATATTCTACACTGGAGCCAGGCTATTAAGTCTGGGAAATTTCAA 2386
 1083 TTGACTGGGGAGGTGAGACCAAAATCTGGAATAATGCAATCAGCCAACTCTCTGTAAG 1142
 2387 TTGACTGGGGAGGTGAGACCAAAATCTGGAATAATGCAATCAGCCAACTCTCTGTAAG 2446
 1143 TACAGAGTCAGAGATATGACGGTCTCTACAGCAATGTGGAGAGGTGAGGACTGGCTT 1202
 2447 TACAATTTAAGGCAATGATGCTCCCAATGCTGATGAGTGTGCTGCTGCTGCTGCTG 2506
 1203 TCAATCCAGAGAGGTGAAATGCTGCTCTGAGGTGACCACTCTCTCTACCAATAG 1262
 2507 GCTGACCTCTGAGGATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 2566
 1263 AATATTCTGATGGGTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1322
 2567 GAAATTCCAAATTAACAATCACTGCACTTTTATCTGCGCAATGATGATGCACTCAAG 2626
 1323 TACATGAAATCACTCTGATGAGGAGGAGAGA 1359
 2627 TACAATGAAATTAATTTCTTGTGAGCAAAAGACAAA 2663

RESULT 9

US-10-043-665B-5
 ; Sequence 5, Application US/10043665B
 ; Patent No. 6582948
 ; GENERAL INFORMATION:
 ; APPLICANT: Bolen, Paul L.
 ; APPLICANT: Cihak, Paul, L.

; APPLICANT: Scharpf Jr., Lewis G.
 ; TITLE OF INVENTION: Recombinant Kid Pregastric Esterase and Methods for Its
 ; FILE OF INVENTION: IFF-0009
 ; FILE REFERENCE: IFF-0009
 ; CURRENT APPLICATION NUMBER: US/10/043,665B
 ; PRIOR FILING DATE: 2002-09-24
 ; PRIOR APPLICATION NUMBER: US 09/186,489
 ; PRIOR FILING DATE: 1998-11-05
 ; NUMBER OF SEQ ID NOS: 7
 ; SOFTWARE: PatentIn Ver. 2.0
 ; SEQ ID NO 5
 ; LENGTH: 8324
 ; TYPE: DNA
 ; ORGANISM: Yeast YE-1 expression vector
 ; US-10-043-665B-5

Query Match 22.0%; Score 449.8; DB 4; Length 8324;
 Best Local Similarity 62.7%; Pred. No. 2.2e-116; Indels 0; Gaps 0;
 Matches 700; Conservative 0; Mismatches 417;

243 GCTGTGGACCCCAAGACGATTCATGAAATATTAGTGAATCATCCAAATCAAGGCTTATCCC 302
 1547 GCTAAGAACCCCTGAAGCAGATGATGATGAGTCAATGATTTCTCTCTGGGGCTTACCA 1606
 303 TGTGAGGATATGATGCTGCAACTGCAAGATGGGTATATCTCTTCTGTAAACAGATTCCT 362
 1607 AGTGAGATGCAATAAAGTTATTAACGCAATGGCTATATCTCTCAAGGCTATCGGATTCCT 1666
 363 CGAGGCTAGTGCACACCTAAGAAAGACAGGTTCCAGGCTCTGGTGTATTCTGCTGAGCATGGC 422
 1667 CATGGAAGATGATGCTAATCATTTAGTCAAGACCTGTTGTTCTGACGATGGT 1726
 423 CTAGTTGAGGTGCTAGCAACTGGAATTTCCAACTGCTGCCCAACATAGCTGGGCTTCATT 482
 1727 CTTCTTGCTCAGCTACAACTGGAATTTCCAACTTCCCAACAGCCTGGGCTTCTCTC 1786
 483 CTGGCAGATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 542
 1787 CTGGCAGATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 1846
 543 AAACACAGACACTCTCCATAGACCAAGATGATGCTGCTGCTGCTGCTGCTGCTGCTG 602
 1847 GAACTTTTACTATTACAGATCTCCCTGATGATTTCAATGATTTCTGCTGCTGCTGCTG 1906
 603 GCTAGTTTGAATTTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 662
 1907 GCTGATATGACTTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 1966
 663 ATCTATTATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 722
 1967 CTACACTATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 2026
 723 CCAGAGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 782
 2027 CCCACACTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 2086
 783 CATGCAAAAGCCCGGAGCAAAATTTTGTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 842
 2087 CACACCCAGAGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 2146
 843 TTGGCAGAAAGAAATTTCTGATCAGATTTGTAGTAAATATCATGTTTACTTCTGGGTG 902
 2147 TTGGTAAACAAATTTCTACCCACACAAATTTTGTGCTGCTGCTGCTGCTGCTGCTG 2206
 903 TGTGGCAGGTGATTTCTGATCAGATTTGTAGTAAATATCATGTTTACTTCTGGGTG 962
 2207 TGCTCTGCTGAGACACTGATGCTCTTTGTAAGATGCTTGTGCTGCTGCTGCTGCTG 2266
 963 AACACACAAATATGAAATGAGCGGAGCAAGTGTATGCTGCCCACTCTTCTGCTGGA 1022
 2267 GACAAATAAAACCTTCAACATGAGTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 2326
 1023 ACATCTGTGCAAAATATTCTACACTGGAGCCAGGCTATTAAGTCTGGGAAATTTCAA 2386
 1083 TTGACTGGGGAGGTGAGACCAAAATCTGGAATAATGCAATCAGCCAACTCTCTGTAAG 1142
 2387 TTGACTGGGGAGGTGAGACCAAAATCTGGAATAATGCAATCAGCCAACTCTCTGTAAG 2446
 1143 TACAGAGTCAGAGATATGACGGTCTCTACAGCAATGTGGAGAGGTGAGGACTGGCTT 1202
 2447 TACAATTTAAGGCAATGATGCTCCCAATGCTGATGAGTGTGCTGCTGCTGCTGCTG 2506
 1203 TCAATCCAGAGAGGTGAAATGCTGCTCTGAGGTGACCACTCTCTCTACCAATAG 1262
 2507 GCTGACCTCTGAGGATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 2566
 1263 AATATTCTGATGGGTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1322
 2567 GAAATTCCAAATTAACAATCACTGCACTTTTATCTGCGCAATGATGATGCACTCAAG 2626
 1323 TACATGAAATCACTCTGATGAGGAGGAGAGA 1359
 2627 TACAATGAAATTAATTTCTTGTGAGCAAAAGACAAA 2663

Db 2327 GCTTCTGTTCAAAACATCTCCATCGGAGACAGGCTATTAACTCTGGGAAATCCAGGT 2386
 QY 1083 TTTGACTGGGGAGTGAGACCAAAATCTGAAAAATGCAATCAGCAACTCTCTGAAGG 1142
 Db 2387 TTTGACTGGGGAGCCCTCAGTTGAGAACCTTAATGCAATTATAATCAGCCCAACCTCCCATC 2446
 QY 1143 TACAGAGTCAGAGATATGACGGTCCCTACAGCAATGTGACAGGAGGTGAGGACTGGGCTT 1202
 Db 2447 TACAAATTAACAGCCATGATGTCCTCAATTCGAGTATGAGTGTGGCCAGACCTGTTG 2506
 QY 1203 TCAATTCAGAGACGCTGAAATGCTGCTCTCTGAGGTGACCAACTCATCTACCAATAG 1262
 Db 2507 GCTGACCCCTCAGGATGTTGACCTTTTGTCTTCAAACTCTCTAATCTCATTCACCAAG 2566
 QY 1263 AATATTCCTGAAATGGGCTCAGCTGGGATTTCACTCTGGGGTTTGGATGCTCTCACCGTATG 1322
 Db 2567 GAATTTCCAAATTAACATCATCTGGACTTTTCTGAGCAATGATGACCTCAGAGTT 2626
 QY 1323 TACAATGAATATCATCTGATGACAGGAGGAGA 1359
 Db 2627 TACAATGAATATTTCTTTTGATGGCAAAAGACAAA 2663

RESULT 10

US-08-227-108-2
 ; Sequence 2, Application US/08227108
 ; Patent No. 5807726
 ; GENERAL INFORMATION:
 ; APPLICANT: Blanchard, Claire
 ; APPLICANT: Benicourt, Claude
 ; APPLICANT: Junien, Jean-Louis
 ; TITLE OF INVENTION: Recombinant Dog Gastric Lipase
 ; NUMBER OF SEQUENCES: 21
 ; CORRESPONDENCE ADDRESS:
 ; ADDRESSEE: Pennie & Edmonds
 ; STREET: 1155 Avenue of the Americas
 ; CITY: New York
 ; STATE: New York
 ; COUNTRY: U.S.A.
 ; ZIP: 10036

COMPUTER READABLE FORM:
 MEDIUM TYPE: Floppy disk
 COMPUTER: IBM PC compatible
 OPERATING SYSTEM: PC-DOS/MS-DOS
 SOFTWARE: PatentIn Release #1.0, Version #1.25
 CURRENT APPLICATION DATA: US/08/227,108
 APPLICATION NUMBER: US/08/227,108
 FILING DATE: 03-APR-1994
 CLASSIFICATION: 435
 ATTORNEY/AGENT INFORMATION:
 NAME: Fanucci, Allan A.
 REGISTRATION NUMBER: 30,256
 REFERENCE/DOCKET NUMBER: 7620-033
 TELECOMMUNICATION INFORMATION:
 TELEPHONE: 212 790-9090
 TELEFAX: 212 869-8864/9741
 TELEX: 66141 PENNIE
 INFORMATION FOR SEQ ID NO: 2:
 SEQUENCE CHARACTERISTICS:
 LENGTH: 1137 base pairs
 TYPE: nucleic acid
 STRANDEDNESS: double
 TOPOLOGY: linear
 MOLECULE TYPE: cDNA to mRNA
 FEATURE:
 NAME/KEY: CDS
 LOCATION: 1..1137
 US-08-227-108-2

Query Match 21.7%; Score 444.2; DB 1; Length 1137;
 Best Local Similarity 62.8%; Pred. No. 3e-117;
 Matches 689; Conservative 0; Mismatches 408; Indels 0; Gaps 0;

QY 250 ACCCAGAGCATTCATGAATATTAGTGAATCATCCAAATCAAGGCTATCCCTGTGAGG 309
 Db 26 ACCTGAAGTGACCATGAATATTAAGTCAGATGATCACTTCTGGGATACCCAGCTGAGG 85
 QY 310 AATATGAAGTCGCAACTGAAGATGGGTATATCTCTTCTCTTAAACAGAGTTCCTCGAGGCC 369
 Db 86 AATATGAAGTTGTGACCGAAGCGGTATATCTCTGGGATCGACAGATTCCTTATGGGA 145
 QY 370 TAGTGCAACCTTAAGAGACAGAGTTCCAGGCTGTGGTGTACTGACGATGGCTAGTTG 429
 Db 146 GGAATAATTCAGAGATATAGGCCGAGACCTTGTGCAATTTTGGCAACAGGTTTCTCTG 205
 QY 430 GAGTGCTAGCAACTGGAATTTCCAACTTCCCAACTAGCTTCCCAACTAGCTTGGCTTCAATTCGCGAG 489
 Db 206 CATCAGCCACAAACTGGATCTCCAACTGCTGCCAAACAAACAGCTTGGCTTCACTCTGGCCG 265
 QY 490 ATGCTGCTTTTGGAGTGGGAAACAGAGGAGGAAACGCTGCTCGAAACACA 549
 Db 266 ACGCCGGGTACGACGTGTGGCTGGGAAACAGAGGAGGAAACACCTGGCCAGGAGGAATC 325
 QY 550 AGACACTCTCCATAGACCAAGATGAGTTCTGGGCTTTTCACTTATGATGAGATGGCTAGGT 609
 Db 326 TGTACTACTCGCCGCACTCCGTCGAATTTCTGGGCTTTTCACTTGGAGATGGCTAAAT 385
 QY 610 TTGACCTTCTGACGTGATAACTTTATTTTGCAGAAACGGGCGCAGGAAAGATCTATT 669
 Db 386 ATGACCTTCCCGCCACCATTTGACTTTCATCTTGAAGAAACGGGACAGGCAAGCTACACT 445
 QY 670 ATGTGCGCTATTACAGGGGCAACCATGGGCTTTATTTGCAATTTTCCACCATGCCAGAGC 729
 Db 446 ACCTTGGCCATTTCCAGGGGCAACCATTTGTTTTCATGCGCTTTTCCACCAATCCCAGC 505
 QY 730 TGGCTCAGAAATCAAAATGATTTTGTCTTTAGACCCATAGCCACTGTTTAAGCATGCA 789
 Db 506 TGGCGAAACGGATCAAAACCTTCTATGCAATAGCTCCCGTTGCCCGCTGGAAGTACACCG 565
 QY 790 AAAGCCCGGAGCCAAATTTTGTGTGCGCAGATATGATGATCAAGGGATTTGTTGGCA 849
 Db 566 AAACCTGTTAAACAACTCATGCTGCTGCTTCTCTTCAAGCTTATATTGGAA 625
 QY 850 AAAAAGAAATTTCTGTATCAGACAGATTTCTCAGACAACTTTGATTTATTTACCTTTGGGCC 909
 Db 626 ACAAATATTTCTACCCACACCACTTCTTTGATCAATTTCTGCGCCACCGAGGTATGCTCC 685
 QY 910 AGGTGATTTCTGATCAGATTTGTAGTAATATATCTTCTCTGCTGGAATTCACACCA 969
 Db 686 GCGAGACGGTGGATCTCTCTGAGCAACGCTCTGTTATCATTTGTGATTTGACACTA 745
 QY 970 ACAATATGAACATGAGCCGAGCAAGTGTATATGCTGCCACACTCTTGTGCGAACTGTG 1029
 Db 746 TGAATTTGAACATGAGTGGCTTGGATGTGTATCTGTACATAATCCAGCAGGAACATCGG 805
 QY 1030 TGCATAATTTCTACACTGGAGCCAGCAGTGAATCTGGTGAATTCGGGCAATTTGACT 1089
 Db 806 TTCAGAACGTCCTCCACTGCTGCGAGCAACGCTCTGTTATCATTTGTGATTTGACACT 865
 QY 1090 GGGGAGGTGAGACCAAAATCTGGAATAATGCAATCAGCCAACTCTCTTAAGGTACAGAG 1149
 Db 866 GGGGAAGCCAGTTTCAGAACATGATGCACTATCATCAGAGCATGCTCCCTACTACAAC 925
 QY 1150 TCAGAGATAGCGGTCCCTACAGCAATGTGGACAGAGGTGAGGATGGCTTTCAATC 1209
 Db 926 TGAAGACATGATGTGCCAATCGCAGTGTGGAAACGGTGGCAACGACTTGTCTGGCGACC 985
 QY 1210 CAGAAGAGCGTGAATAATGCTGCTCTCTGAGGTGACCAACCTCATCTACCAAGAAATTC 1269
 Db 986 CTCAGATGTTGACCTTTTCTTCCAGCTCCCACTCTCAATTTACCAAGAGATTC 1045
 QY 1270 CTGAATGGGCTCAGCTGGAATTTCACTGGGGTTTGGATGCTCTCAGCGTATGCAATG 1329
 Db 1046 CTCCTTCAATCACTTGGACTTTATCTGGGCCCATGGATGCCCTCAAGCGGTTTACAATG 1105

QY 1330 AAATCATCCATCTGATG 1346
 Db 1106 AAATGTTTCCATGATG 1122

RESULT 11

US-09-073-674-2
 ; Sequence 2, Application US/09073674
 ; Patent No. 5998189

; GENERAL INFORMATION:
 ; APPLICANT: Blanchard, Claire
 ; APPLICANT: Benicourt, Claude
 ; APPLICANT: Junien, Jean-Louis
 ; TITLE OF INVENTION: Recombinant Dog Gastric Lipase
 ; NUMBER OF SEQUENCES: 21
 ; CORRESPONDENCE ADDRESS:
 ; ADDRESSEE: Warner-Lambert Company
 ; STREET: 2800 Plymouth Road
 ; CITY: Ann Arbor
 ; STATE: Michigan
 ; COUNTRY: U.S.A.
 ; ZIP: 48105

; COMPUTER READABLE FORM:
 ; MEDIUM TYPE: Floppy disk
 ; COMPUTER: IBM PC compatible
 ; OPERATING SYSTEM: PC-DOS/MS-DOS
 ; SOFTWARE: Patent In Release #1.0, Version #1.25
 ; CURRENT APPLICATION DATA:
 ; APPLICATION NUMBER: US/09/073,674
 ; FILING DATE:
 ; CLASSIFICATION:

; ATTORNEY/AGENT INFORMATION:
 ; NAME: Crissey, Todd M.
 ; REGISTRATION NUMBER: 37,807
 ; REFERENCE/DOCKET NUMBER: 5072-DI-66-TWC
 ; TELECOMMUNICATION INFORMATION:
 ; TELEPHONE: 734 622-7530
 ; TELEFAX: 734 622-1553
 ; TELEX:
 ; INFORMATION FOR SEQ ID NO: 2:

; SEQUENCE CHARACTERISTICS:
 ; LENGTH: 1137 base pairs
 ; TYPE: nucleic acid
 ; STRANDEDNESS: double
 ; TOPOLOGY: linear
 ; MOLECULE TYPE: cDNA to mRNA
 ; FEATURE:
 ; NAME/KEY: CDS
 ; LOCATION: 1..1137
 ; US-09-073-674-2

Query Match 21.7%; Score 444.2; DB 2; Length 1137;
 Best Local Similarity 62.8%; Pred. No. 3e-117;
 Matches 689; Conservative 0; Mismatches 408; Indels 0; Gaps 0;

QY 250 ACCGAGAGCAATCATGAATATTAGTGAATCATCAACATCAAGCATCCCTGAGG 309
 Db 25 ACCCTGAGTGACCATGATATAGTCAGATGATCACTACTGGGGATACCCAGCTGAG 85
 QY 310 AATATGAATGCGCACTGAAGATGGGTATATCTTCTGTTAAACAGGATTCCTCGAGGCC 369
 Db 86 AATATGAATGTTGACCGAAGACGGTTATATCTTGGGATCGACAGAAATTCCTTATGGGA 145
 QY 370 TAGTGAACCTAAGAGACAGGTTCCAGGCGTGTGTACTGTGATGCGATGCGCTAGTTG 429
 Db 146 GGNAAATTCAGAGATATAGCCGGAGACCTTGTGATTTTGCAACAGGTTTGTCTG 205
 QY 430 GAGGTCTAGCACTGGATTTTCAACCTGCCCAACATAGCTGGGCTTCATCTTGGCAG 489
 Db 206 CATCAGCCCAAACTGGATCTCAACCTGCCCAACAGAGCTGCGCTTCATCTTGGCAG 265
 QY 490 ATGCTGTTTACGCTGTGGATGGGGAACAGCAGGGAACCCCTGCTCGAAGAACACA 549

Db 266 ACGCCGGGTACGACGTGTGCTGGGGAACAGCAGGGGCAACACCTGGGCCAGAGGAATC 325
 QY 550 AGACACTCTCCATAGACCAAGATGAGTTCTGGGCTTTTCAGTTATGATGAGATGGCTAGGT 609
 Db 326 TGTACTACTCGCCGACTCGTGAATTCCTGGGCTTTTCAGCTTTGACGAGATGGCTAAAT 385
 QY 610 TTGACCTTCTGCACTGATGAATTTATTTTTCAGAAAACGGGCCAGGAAAAGATCTATT 669
 Db 386 ATGACCTTCCCGCCACCAATTCATCTTTGAAGAAAACGGGACAGGCAAGCTACACT 445
 QY 670 ATGTGCGCTATTACAGGGGCAACCAATGGGCTTTATTTGCAATTTCCACCATCCAGAGC 729
 Db 446 ACCTTGGCAATTCAGGGGCAACCAATGGGCTTTTCATGCGCTTTCCACCAATCCCAAGC 505
 QY 730 TGGCTCAGAAAATCAAAATGTAATTTGCTTTAGCACCCATAGCCATGTTAAGCATGCAA 789
 Db 506 TGGCGAAACGGATCAAAACCTTCTATGCAATAGCTCCCGTTGCCCGTGAAGTACACCG 565
 QY 790 AAAGCCCGGGAGCAAAATTTTGTGCTGCGAGATATGATCAAGGGATTTGTTGGCA 849
 Db 566 AAACCTGTTAAACAAACTCATGCTGCTCCCTTCCTCTCAAGCTTATATTTGGAA 625
 QY 850 AAAAAGAAATTCGTATCAGACCAAGATTTCTCAGCAACTGTGTATTTACCTTTGTGGCC 909
 Db 626 ACAAATATTTCTACCCACACCACTTCTTTGATCAATTTCTGCGCCACCGAGGTATGCTCC 685
 QY 910 AGGTGATTTCTGATCAGATTTGTAATATATCATGTTACTTCTGGGTGGATTCACACCA 969
 Db 686 CGGAGCGGTGGATCTCTCTGAGCAACGCGCTGTTATCATTTGTGATTTGACACTA 745
 QY 970 ACAATATGAACATGAGCCGAGCAAGTGTATATGCTGCCCACTCTTGTGGAACATCTG 1029
 Db 746 TGAACCTGAAACATGAGTGGCTTGGATGTGTATCTGTCAATAATCCAGCAGGAACATCG 805
 QY 1030 TGCAAAATATTTCTACACTGAGCCAGCAGTGAATTTCTGGTGAATCCCGGCAATTTGACT 1089
 Db 806 TTCAAGACGTGCTCCCACTGGTCCCGCTGTTAAGTCTGGGAAGTTCCCAAGCTTTTACT 865
 QY 1090 GGGGAGTGAGACCAAAATCTCGAAAATGCAATCAGCCAACTCCTCTAAGGTACAGAG 1149
 Db 866 GGGGAGCCCGAGTTCAAGACATGATGCACTATCATCAGAGCATGCTCCTCTACTACAAC 925
 QY 1150 TCAGAGATGACGGTCCCTCAGCAATGTGAGCAGAGGTGAGGATGGCTTCTTCAATC 1209
 Db 926 TGACAGACATGATGTCGAATCGCAATCGCAGTGTGGAAACGGTGCACAGCTTCTGCGCC 985
 QY 1210 CAGAGAGCTGAAAATGCTGCTCTCTGAGGTGACCAACCTCATCTACCAATAAGAAATTC 1269
 Db 986 CTCAGATGTTGACCTTTTGTCTTCCAGCTCCCAATCTCATTTACCAAGAGATTC 1045
 QY 1270 CTGAATGGGCTCAGGTGATTTTCATCTGGGTTTGGATGCTCTCCCGTATGTACAATG 1329
 Db 1046 CTCCTTACATCACTTGGACTTTATCTGGGCCATGGATGCCCTCAAGCGGTTTACAATG 1105
 QY 1330 AAATCATCCATCTGATG 1346
 Db 1106 AAATGTTTCCATGATG 1122

RESULT 12

US-08-227-108-4
 ; Sequence 4, Application US/08227108
 ; Patent No. 5807726
 ; GENERAL INFORMATION:
 ; APPLICANT: Blanchard, Claire
 ; APPLICANT: Benicourt, Claude
 ; APPLICANT: Junien, Jean-Louis
 ; TITLE OF INVENTION: Recombinant Dog Gastric Lipase
 ; NUMBER OF SEQUENCES: 21
 ; CORRESPONDENCE ADDRESS:
 ; ADDRESSEE: Pennie & Edmonds
 ; STREET: 1155 Avenue of the Americas
 ; CITY: New York

[illegible]

```

RESULT 13
US-09-073-674-4
; Sequence 4, Application US/09073674
; Patent No. 5998189
; GENERAL INFORMATION:
; APPLICANT: Blanchard, Claire
; APPLICANT: Benicourt, Claude
; APPLICANT: Junien, Jean-Louis
; TITLE OF INVENTION: Recombinant Dog Gastric Lipase
; NUMBER OF SEQUENCES: 21
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Warner-Lambert Company
; STREET: 2800 Plymouth Road
; CITY: Ann Arbor
; STATE: Michigan
; COUNTRY: U.S.A.
; ZIP: 48105
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patcin In Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/09/073,674
; FILING DATE:
; CLASSIFICATION:
; ATTORNEY/AGENT INFORMATION:
; NAME: Crissey, Todd M.
; REGISTRATION NUMBER: 37,807
; REFERENCE/DOCKET NUMBER: 5072-D1-66-TWC
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 734 622-7530
; TELEFAX: 734 622-1553

```

Qy	1090	GGGGAGTGAGACCAAAAATCTGGA AAAATGCAATCAGCCAACTCCTGTAAAGTACAGAG	1149
Db	869	GGGGAAGCCAGGTTTCAGAAATGATGCACTATCATCAGAGCATGCGCTCCCTACTACAACC	928
Qy	1150	TCAGAGATATGAGCGTCCCTACAGCAATGTGGACAGGAGTCAGGACTGGCTTTCAAATC	1209
Db	929	TGACAGACATGATGTGCCAATCGCAAGTGTGGAACGGTGGCAACGACTTGTGCGCGACC	988
Qy	1210	CAGAAGACGTGAAAAATGCTGCTCTCTGAGGTAGCAACACCTCATCTACCAATAAATATTC	1269
Db	989	CTCAGCATGTGACCTTTTGGTTTCCAAAGCTCCCAATCTCATTTATACCAAGGAAGATTC	1048
Qy	1270	CTGAATGGGCTCAGCGGATTTTCATCTGGGGTTTGGATGCTCCTCCCGTATCTGCAATG	1329
Db	1049	CTCCTTACATCACTTGGACTTTATCTGGGCCATGGATGCCCTCAAGCGTTTACAATG	1108
Qy	1330	AAATCATCCATCTGATG 1346	
Db	1109	AAATTGTTTCCATGATG 1125	
RESULT 14			
US-08-227-108-6			
; Sequence 6, Application US/08227108			
; Patent No. 5807726			
GENERAL INFORMATION:			
; APPLICANT: Blanchard, Claire			
; APPLICANT: Benicourt, Claude			
; APPLICANT: Junien, Jean-Louis			
; TITLE OF INVENTION: Recombinant Dog Gastric Lipase			
; NUMBER OF SEQUENCES: 21			
; CORRESPONDENCE ADDRESS:			
; ADDRESSEE: Pennie & Edmonds			
; STREET: 1155 Avenue of the Americas			
; CITY: New York			
; STATE: New York			
; COUNTRY: U.S.A.			
; ZIP: 10036			
COMPUTER READABLE FORM:			
; MEDIUM TYPE: Floppy disk			
; COMPUTER: IBM PC compatible			
; OPERATING SYSTEM: PC-DOS/MS-DOS			
; SOFTWARE: Patentin Release #1.0, Version #1.25			
CURRENT APPLICATION DATA:			
; APPLICATION NUMBER: US/08/227,108			
; FILING DATE: 03-APR-1994			
CLASSIFICATION: 435			
ATTORNEY/AGENT INFORMATION:			
; NAME: Fanucci, Allan A.			
; REGISTRATION NUMBER: 30,256			
; REFERENCE/DOCKET NUMBER: 7620-033			
TELECOMMUNICATION INFORMATION:			
; TELEPHONE: 212 790-9090			
; TELEFAX: 212 869-8864/9741			
; TELEX: 66141 PENNIE			
INFORMATION FOR SEQ ID NO: 6:			
SEQUENCE CHARACTERISTICS:			
; LENGTH: 1146 base pairs			
; TYPE: nucleic acid			
; STRANDEDNESS: double			
; TOPOLOGY: linear			
; MOLECULE TYPE: cDNA to mRNA			
US-08-227-108-6			
Query Match 21.7%; Score 444.2; DB 1; Length 1146;			
Best Local Similarity 62.8%; Pred. No. 3e-117;			
Matches 689; Conservative 0; Mismatches 408; Indels 0; Gaps 0;			
Qy	250	ACCAGAAGCATTCATGAATATTAGTGAATCATCAACATCAAGCTATCCTGTGTGAGG	309
Db	26	ACCTTGAAGTGACCATGAATAAAGTCAGATGATCACTCTGGGATACCCAGCTGAGG	95
Qy	310	AATPATGAAGTCGCAACTGAAGATGGGTATATCTTTCTGTTAACAGGATTTCTCGAGGCC	369

Db 86 AATATGAAGTTGTACGACGAGTATATATCTTGGATCGACAGATTCCTTATGGGA 145
 Qy 370 TAGTGCAACCTAAGAGACAGGTTCCAGGCTGTGGTGTACTGACAGATGGCTAGTTG 429
 Db 146 GGAATAATTCAGAGATATAGCGCGAGACCTGTTCATTTTGGCAACGGTTTGTCTG 205
 Qy 430 GAGGTGTAGCAACTGGATTTCACACCTGCCCAATAGCTGGGCTTCATTTCTGCGAG 489
 Db 206 CATCAGCAAACTGGATCTCAAACCTGCCCAACACAGCTGGCTTCATCTGGCGG 265
 Qy 490 ATGCTGGTTTGAAGTGTGATGGGAAACAGCAGGGGAAACCGCTGTCTCGAAACACA 549
 Db 266 ACGCGGGTACAGCTGTGGCTGGGAAACAGCAGGGGCAACACCTGGCGCAGGAGAA 325
 Qy 550 AGACACTCTCCATAGACCAAGATGTTCTGGGCTTTTCAGTTATGATGAGATGGTAGGT 609
 Db 326 TGTACTACTCCCGCAGCTCCCTGAAATCTGGGCTTTTCAGTTTGCAGATGGCTAAT 385
 Qy 610 TTGACCTTCTGCAAGTATAAATTTTTCAGAAACCGGCGCAGGAAAGATCTATT 669
 Db 386 ATGACCTTCCCGCAGCTTACATTTCACTTTGAAGAAACCGGACAGCAAGCTACACT 445
 Qy 670 ATGCTGGCTATTCAGGGCACCACCTGGCTTTTTCAGTTTTCACCATGCCAGAGC 729
 Db 446 ACGTTGGCTATTCAGGGCACCACCTGGCTTTTTCAGTTTTCACCATGCCAGAGC 505
 Qy 730 TGGCTCAGAAATCAAAATGTTATTTGCTTTAGCACCCTAGCCACTGTTAAGCATGCAA 789
 Db 506 TGGGCAACCGATCAAAACCTTCTATGATAGCTCCGTTGCCACCGTGAAGTACACCG 565
 Qy 790 AAAGCCCGGACCAAAATTTTGTGTCGCCAGATATGATGATCAAGGATTTGTGGCA 849
 Db 566 AAACCTGTTAAACAACTCATGCTGCTGCTTCTGCTTCTTCAAGCTTATATTGGAA 625
 Qy 850 AAAAGAAATTTCTGTATCAGACCAAGATTTCTCAGACAACTTGTATTATTCCTTTGGCC 909
 Db 626 ACATAATTTTACCCACACCACTTCTTTGATCAATTTCTGCCACCGAGTATGCTCCC 685
 Qy 910 AGGTGATTTTGAATCAGATTTGTAGTATATATCATGTTACTTCTGGTGGATTCACACA 969
 Db 686 GCGAGACGGTGGATCTCTCTGACGCAACGGCTGTTTATCATTTGTGATTTGACACTA 745
 Qy 970 ACAATATGAATGAGCGGACGAGTATGATGCTGCCACACTTCTGCTGGAACTCTG 1029
 Db 746 TGAATTTGAATGAGTCTGCTGGATGTTATCTCTACATAATCCAGCAGGACATCGG 805
 Qy 1030 TGCAAAATATTTCACTAGGAGCGGAGGAGTGAATTTCTGGTGAATCTGGGCACTTGA 1089
 Db 806 TTCAGACGTGCTCCACTGGTCCGAGGCTGTTAAGTCTGGGAAGTTCCAACTTTTGA 865
 Qy 1090 GGGGAGTGAGACCAAAATCTGGAATAATGCAATTCAGCCAACTCTGTAAGGTACAGAG 1149
 Db 866 GGGGAGGCCAGTTTCAGAACATGATGCACTATCATCAGAGATGGCTCCCTACTACACC 925
 Qy 1150 TCAGAGATATAGCGTCCCTACAGCAATGTGACAGGAGGTGAGAGCTGGCTTTCAAATC 1209
 Db 926 TCACAGACATGATGTGCAATTCGCAATGTGGAACGGTGGCAACGACTTGTGGCGGACC 985
 Qy 1210 CAGAAAGCTGGAATGCTGCTCTGAGGTCACCAACCTCATCTACCATGAAGATATTC 1269
 Db 986 CTCACAGATTTGACCTTTGCTTTTCCAAAGCTCCCAATCTCATTTACCAAGGAGATTC 1045
 Qy 1270 CTGAATGGGCTCACGTGATTTTCACTGGGGTTTGGATGCTCTCACCGTATGTACAA 1329
 Db 1046 CTCCTTACAATCACTTGGACTTTATCTGGGCAATGATGCTCCCTCAAGCGGTTTACA 1105
 Qy 1330 AATCATCATCTGATG 1346
 Db 1106 AATTTGTTTCCATGATG 1122

RESULT 15

US-09-073-674-6
 ; Sequence 6, Application US/09073674
 ; Patent No. 5998189
 ; GENERAL INFORMATION:
 ; APPLICANT: Blanchard, Claire
 ; APPLICANT: Benicourt, Claude
 ; APPLICANT: Junien, Jean-Louis
 ; TITLE OF INVENTION: Recombinant Dog Gastric Lipase
 ; NUMBER OF SEQUENCES: 21
 ; CORRESPONDENCE ADDRESS:
 ; ADDRESSEE: Warner-Lambert Company
 ; STREET: 2800 Plymouth Road
 ; CITY: Ann Arbor
 ; STATE: Michigan
 ; COUNTRY: U.S.A.
 ; ZIP: 48105
 ; COMPUTER READABLE FORM:
 ; MEDIUM TYPE: Floppy disk
 ; COMPUTER: IBM PC compatible
 ; OPERATING SYSTEM: PC-DOS/MS-DOS
 ; SOFTWARE: PatentIn Release #1.0, Version #1.25
 ; CURRENT APPLICATION DATA:
 ; APPLICATION NUMBER: US/09/073,674
 ; FILING DATE:
 ; CLASSIFICATION:
 ; ATTORNEY/AGENT INFORMATION:
 ; NAME: Crissey, Todd M.
 ; REGISTRATION NUMBER: 37,807
 ; REFERENCE/DOCKET NUMBER: 5072-D1-66-TWC
 ; TELECOMMUNICATION INFORMATION:
 ; TELEPHONE: 734 622-7530
 ; TELEFAX: 734 622-1553
 ; TELEX:
 ; INFORMATION FOR SEQ ID NO: 6:
 ; SEQUENCE CHARACTERISTICS:
 ; LENGTH: 1146 base pairs
 ; TYPE: nucleic acid
 ; STRANDEDNESS: double
 ; TOPOLOGY: linear
 ; MOLECULE TYPE: cDNA to mRNA
 ; US-09-073-674-6

Query Match 21.7%; Score 444.2; DB 2; Length 1146;
 Best Local Similarity 62.8%; Pred. No. 3e-117;
 Matches 689; Conservative 0; Mismatches 408; Indels 0; Gaps 0;
 Qy 250 ACCGAGACCATTCATGATATATAGTGAATCATCATCAACATCAAGCTATCCCTGTGAGG 309
 Db 26 ACCCTGAAGTACCATGAATATAGTCAGATGATCACTCTGGGATACCCAGCTGAGG 85
 Qy 310 AATATGAAGTCTGCAACTGGATTTCCAACTGCCCAACATAGCTGGGCTTCATCTTGGCAG 489
 Db 86 AATATGAAGTCTGACCGAAGACGCTTATATCTTGGGATCGACAGAAATTCCTTATGGGA 145
 Qy 370 TAGTGCAACTAAGAGACAGGTTCCAGGCTGTGGTGTACTGACAGATGGCTAGTTG 429
 Db 146 GGAATAATTCAGAGATATAGCGCGAGACCTGTTCATTTTGGCAACGGTTTGTCTG 205
 Qy 430 GAGGTGTAGCAACTGGATTTCCAACTGCCCAACATAGCTGGGCTTCATCTTGGCAG 489
 Db 206 CATCAGCAAACTGGATTCCTCAACCTGCCCAACACAGCTGGCTTCATCTGGCGG 265
 Qy 490 ATGCTGGTTTGAAGTGTGATGGGAAACAGCAGGGGAAACCGCTGTCTCGAAACACA 549
 Db 266 ACGCGGGTACAGCTGTGGCTGGGAAACAGCAGGGGCAACACCTGGCGCAGGAGAA 325
 Qy 550 AGACACTCTCCATAGACCAAGATGTTCTGGGCTTTTCAGTTATGATGAGATGGTAGGT 609
 Db 326 TGTACTACTCCCGCAGCTCCCTGAAATCTGGGCTTTTCAGTTTGCAGATGGCTAAT 385
 Qy 610 TTGACCTTCTGCAAGTATAAATTTTTCAGAAACCGGCGCAGGAAAGATCTATT 669
 Db 386 ATGACCTTCCCGCAGCTTACATTTCACTTTGAAGAAACCGGACAGCAAGCTACACT 445

Qy	670	ATGTCGGCTATTACAGGGACCAACCATGGGCTTTATTGTCATTTTCCACCATGCCAGGC	729
Db	446	ACGTTGGCCATTCAGGGACCAACCATGGGCTTTATTGTCATTTTCCACCATGCCAGGC	505
Qy	730	TGGCTCAGAAATCAGAAATGATTTTCTTATGACCCATAGCCACTGTTAAGCATGCAA	789
Db	506	TGGGAAACGGATCAAAACCTTCTATGATAGCTCCGTTGCCACCGTGAAGTACACCG	565
Qy	790	AAAGCCCGGACCAAAATTTTGTGTCGTCAGATATGATGATCAAGGATGTTGGCA	849
Db	566	AAACCTTGTAAACAACTATGCTGCTCCCTTCTTCAAGCTTATATTGGAA	625
Qy	850	AAAAAGAAATTTCTATCAGACCAAGATTTCTCAGACAACTGTTATTATTTTGTGGCC	909
Db	626	ACAAATATTTCTACCCACACCACTTCTTGTATCAATTTCTGGCCACCGAGGTATGCTCC	685
Qy	910	AGGTGATTTCTGATCAGATTTGTAGTAATATCATGTTACTTCTGGTGGATTTCAACCA	969
Db	686	CGGAGCGGTGATCTCTCTGACGAAACGCGCTGTTTATCTTTGGAATTTGACACTA	745
Qy	970	ACAATATGAACATGAGCGGAGCAAGTGTATATGTCGCCACACTTGTGCGAACATCTG	1029
Db	746	TGAATTTGAACATGAGTGGCTTGGATCTGTATCTGTACATATCCAGCAGGAACATCG	805
Qy	1030	TGCAAAATATTTCTACACTGAGCCAGCCAGTGAATTTCTGTGAACTCCGGGCAATTTGACT	1089
Db	806	TTCAAGACGTGCTCCACTGCTGCCAGGCTGTTAAGTCTGGAAAGTTCCAAGCTTTTACT	865
Qy	1090	GGGGAGTGAAGACCAAAATCTGGAAAATGCAATCAGCCAACTCTCTGTAAGGTACAGAG	1149
Db	866	GGGAAGCCAGTTCAAGACATGATGCACTATCATCAGAGCATGCTCCCTACTACACC	925
Qy	1150	TCAGAGATATGACGGTCCCTTACAGCAATGTGGACAGAGGTGAGGCTGAGGCTTTCAAATC	1209
Db	926	TGACAGACATGCAATGTGCCAATCGCAGTGTGGAAACGGTGGCAACGACTTGTGGCGGAC	985
Qy	1210	CAGAAGACGTGAAATGCTCTCTCAGGTGACCACTCTACCATTAAGAAATATTC	1269
Db	986	CTCAGATGTGACCTTTTGGTTTCCAAAGCTCCCAATCTCATTTACACAGGAATTC	1045
Qy	1270	CTGAATGGGCTCAGTGGATTTTCAATGTTGGGTTGGATGCTCTCACCCTATGTACAATG	1329
Db	1046	CTCCTTACAATCACTTGGACTTTTATCTGGGCCATGGATGCCCTCAAGCGGTTTACAATG	1105
Qy	1330	AAATCATCCATCTGATG	1346
Db	1106	AAATGTTCCATGATG	1122

Search completed: February 19, 2004, 10:33:53
Job time : 135.478 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2004 Compugen Ltd.

OM nucleic - nucleic search, using sw model

Run on: February 19, 2004, 08:29:29 ; Search time 739.122 Seconds
(without alignments)
10186.874 Million cell updates/sec

Title: US-10-042-431-45
Perfect score: 2044
Sequence: 1 gtcgaccacgctccggg.....aaaaaaaggcgccgc 2044

Scoring table: IDENTITY NUC
Gapop 10.0, Gapext 1.0

Searched: 2449703 seqs, 1841816367 residues
Total number of hits satisfying chosen parameters: 4899406

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : Published Applications NA:
1: /cgn2_6/ptodata/1/pubna/US07_PUBCOMB.seq.*
2: /cgn2_6/ptodata/1/pubna/PCT_NEW_PUB.seq.*
3: /cgn2_6/ptodata/1/pubna/US06_NEW_PUB.seq.*
4: /cgn2_6/ptodata/1/pubna/US06_PUBCOMB.seq.*
5: /cgn2_6/ptodata/1/pubna/US07_NEW_PUB.seq.*
6: /cgn2_6/ptodata/1/pubna/PCTUS_PUBCOMB.seq.*
7: /cgn2_6/ptodata/1/pubna/US08_NEW_PUB.seq.*
8: /cgn2_6/ptodata/1/pubna/US08_PUBCOMB.seq.*
9: /cgn2_6/ptodata/1/pubna/US09_PUBCOMB.seq.*
10: /cgn2_6/ptodata/1/pubna/US09_PUBCOMB.seq.*
11: /cgn2_6/ptodata/1/pubna/US09_PUBCOMB.seq.*
12: /cgn2_6/ptodata/1/pubna/US09_NEW_PUB.seq.*
13: /cgn2_6/ptodata/1/pubna/US09_NEW_PUB.seq.*
14: /cgn2_6/ptodata/1/pubna/US10_PUBCOMB.seq.*
15: /cgn2_6/ptodata/1/pubna/US10_PUBCOMB.seq.*
16: /cgn2_6/ptodata/1/pubna/US10_NEW_PUB.seq.*
17: /cgn2_6/ptodata/1/pubna/US60_NEW_PUB.seq.*
18: /cgn2_6/ptodata/1/pubna/US60_PUBCOMB.seq.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	2044	100.0	2044	11	US-09-759-130B-415
2	2044	100.0	2044	14	US-10-042-431-45
3	1993.8	97.5	2035	10	US-09-835-996A-20
4	1343.4	65.7	1384	10	US-09-835-996A-9
5	1269	62.1	1269	11	US-09-759-130B-416
6	1269	62.1	1269	14	US-10-042-431-46
7	1126	55.1	1126	10	US-09-811-825-3
8	890.4	43.6	18554	10	US-09-811-825-3
9	527.8	25.8	2481	10	US-09-880-107-3878
10	527.8	25.8	2858	15	US-10-198-846-13431
11	527.8	25.8	6952	13	US-10-133-013-144
12	525.4	25.7	2853	13	US-09-971-392-116
13	462.2	22.6	2806	12	US-10-108-260A-1262
14	456.8	22.3	1411	12	US-10-420-564-3
15	454.2	22.2	1194	13	US-10-312-088-7

16	453.6	22.2	1134	12	US-10-420-564-4
17	450.8	22.1	1194	14	US-10-056-744B-3
18	450.8	22.1	1325	14	US-10-056-744B-1
19	450.4	22.0	1134	12	US-10-420-564-1
20	449.8	22.0	8324	12	US-10-420-564-5
21	443.6	21.7	1200	13	US-10-221-097-9
22	424.8	20.8	1260	13	US-10-015-115-31
23	414.2	20.3	1360	14	US-10-003-302-1
24	346.4	16.9	1267	12	US-10-051-874-27
25	344.8	16.9	1267	12	US-10-051-874-29
26	332.8	16.3	1098	13	US-10-312-088-6
27	327.8	16.0	1195	12	US-10-051-874-31
28	303.6	14.9	634	15	US-10-198-846-9462
29	300.4	14.7	1377	13	US-10-015-115-29
30	196.6	9.6	495	15	US-10-060-036-4296
31	192.8	9.4	839	15	US-10-198-846-6768
32	171.2	8.4	534	15	US-10-060-036-4334
33	165.4	8.1	510	15	US-10-060-036-4326
34	165.4	8.1	522	15	US-10-060-036-4325
35	161.8	7.9	635	13	US-10-099-926-1834
36	161.8	7.9	635	14	US-10-033-528-1834
37	152.6	7.5	460	12	US-10-062-674-253
38	128.4	6.3	362	10	US-09-960-352-13032
39	124.4	6.1	624	13	US-10-027-632-214253
40	124.4	6.1	624	14	US-10-027-632-214253
41	122.6	6.0	419	11	US-09-918-995-4162
42	113	5.5	1242	12	US-10-260-238-704
43	104.2	5.1	22067	14	US-10-003-302-3
44	100.4	4.9	371	11	US-09-918-995-37615
45	89.4	4.4	367	10	US-09-960-352-13508

ALIGNMENTS

RESULT 1

US-09-759-130B-415
Sequence 415, Application US/09759130B
Publication No. US2003002279A1
GENERAL INFORMATION:
APPLICANT: Millennium Pharmaceuticals, Inc.
APPLICANT: McCarthy, Sean A
APPLICANT: Fraser, Christopher C
APPLICANT: Sharp, John D
APPLICANT: Barnes, Thomas S
APPLICANT: Kirst, Susan J
APPLICANT: Mackay, Charles R
APPLICANT: Myers, Paul S
APPLICANT: Leiby, Kevin R
APPLICANT: Wrighton, Nicolas
APPLICANT: Goodearl, Andrew
APPLICANT: Holtzman, Douglas A
TITLE OF INVENTION: NOVEL GENES ENCODING PROTEINS HAVING
TITLE OF INVENTION: PROGNOSTIC, DIAGNOSTIC, PREVENTIVE, THERAPEUTIC, AND OTHER
TITLE OF INVENTION: USES
FILE REFERENCE: MP100-5350NMIM
CURRENT APPLICATION NUMBER: US/09/759,130B
CURRENT FILING DATE: 2002-09-16
PRIOR FILING DATE: US 09/479,249
PRIOR FILING DATE: 2000-01-07
PRIOR APPLICATION NUMBER: US 09/559,497
PRIOR FILING DATE: 2000-04-27
PRIOR APPLICATION NUMBER: US 09/578,063
PRIOR FILING DATE: 2000-05-24
PRIOR APPLICATION NUMBER: US 09/333,159
PRIOR FILING DATE: 1999-06-14
PRIOR APPLICATION NUMBER: US 09/596,194
PRIOR FILING DATE: 2000-07-14
PRIOR APPLICATION NUMBER: US 09/342,364
PRIOR FILING DATE: 1999-06-29
PRIOR APPLICATION NUMBER: US 09/608,452
PRIOR FILING DATE: 2000-06-30
PRIOR APPLICATION NUMBER: US 09/393,996

```

; PRIOR FILING DATE: 1999-09-10
; PRIOR APPLICATION NUMBER: US 09/602,871
; PRIOR FILING DATE: 2000-06-23
; PRIOR APPLICATION NUMBER: US 09/420,707
; PRIOR FILING DATE: 1999-10-19
; NUMBER OF SEQ ID NOS: 460
; SOFTWARE: Fast-Seq for Windows Version 4.0
; SEQ ID NO 415
; LENGTH: 2044
; TYPE: DNA
; ORGANISM: Homo sapiens
; US-09-759-1308-415

Query Match      100.0%; Score 2044; DB 11; Length 2044;
Best Local Similarity 100.0%; Pred. No. 0;
Matches 2044; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 GTCGACCCAGCGTCGCGGGAATTCGAGCAGGGAATATCTGAAGAGTTTAAACCCAC 60
Db 1 GTCGACCCAGCGTCGCGGGAATTCGAGCAGGGAATATCTGAAGAGTTTAAACCCAC 60

QY 61 AAATTCCTCTTACTTTAGAAATAGTTGTTACATTCGCGAGGAAATAAATGAGATGTT 120
Db 61 AAATTCCTCTTACTTTAGAAATAGTTGTTACATTCGCGAGGAAATAAATGAGATGTT 120

QY 121 GACCATGTTGGAAACCTCTCAAGACAGTGGATGTTCTCAGACAGAAATGGAATGTGC 180
Db 121 GACCATGTTGGAAACCTCTCAAGACAGTGGATGTTCTCAGACAGAAATGGAATGTGC 180

QY 181 TTCTGATTCCTGGTGGCGTATGTTCCAGAGAAATGTGAATTCAGTACATATGCCAACTA 240
Db 181 TTCTGATTCCTGGTGGCGTATGTTCCAGAGAAATGTGAATTCAGTACATATGCCAACTA 240

QY 241 AAGCTGTGACCCGAGAGCAATTCATGAAATATGTAATTCAGTACATATGCCAACTA 300
Db 241 AAGCTGTGACCCGAGAGCAATTCATGAAATATGTAATTCAGTACATATGCCAACTA 300

QY 301 CCTGTGAGGAATATGAAGTGCACCTCAAGATGGGTATATCTCTTCTGTTAAACAGGATC 360
Db 301 CCTGTGAGGAATATGAAGTGCACCTCAAGATGGGTATATCTCTTCTGTTAAACAGGATC 360

QY 361 CTCGAGGCTAGTGCAACCTAAGAGACAGGTTCCAGGCGTGTGGTGTACTGCAGCATG 420
Db 361 CTCGAGGCTAGTGCAACCTAAGAGACAGGTTCCAGGCGTGTGGTGTACTGCAGCATG 420

QY 421 GCCTAGTGTGAGGCTAGCACTGCAATTCGCAATTCGCAATTCGCAATTCGCAATTCGCA 480
Db 421 GCCTAGTGTGAGGCTAGCACTGCAATTCGCAATTCGCAATTCGCAATTCGCAATTCGCA 480

QY 481 TTCTGCGAGATGCTGGTGTGAGCTGTGGATGGGAAACAGCAGGGGAAACGCTGTGCTC 540
Db 481 TTCTGCGAGATGCTGGTGTGAGCTGTGGATGGGAAACAGCAGGGGAAACGCTGTGCTC 540

QY 541 GAAACACAGACACTCTCCATAGACCAAGTGTGCTGGGCTTTTCAGTTATGATGAGA 600
Db 541 GAAACACAGACACTCTCCATAGACCAAGTGTGCTGGGCTTTTCAGTTATGATGAGA 600

QY 601 TGGCTAGGTTGACCTTCCTGCAATGATAAATTTATTTTCAGAGAAACGGGCGAGGAAA 660
Db 601 TGGCTAGGTTGACCTTCCTGCAATGATAAATTTATTTTCAGAGAAACGGGCGAGGAAA 660

QY 661 AGATCTATATGCTGGCTATTCAGAGGACCAACCATGGGCTTTATGCAATTTTCACCA 720
Db 661 AGATCTATATGCTGGCTATTCAGAGGACCAACCATGGGCTTTATGCAATTTTCACCA 720

QY 721 TGCCAGAGCTGGCTCAGAAATCAAAATGATTTTGTGCTTTAGACCCATAGCCACTGTTA 780
Db 721 TGCCAGAGCTGGCTCAGAAATCAAAATGATTTTGTGCTTTAGACCCATAGCCACTGTTA 780

QY 781 AGCATGCAAAAGCCCGGAGCCAAATTTTGTGCTGCCAGATGATGATCAAGGAT 840
Db 781 AGCATGCAAAAGCCCGGAGCCAAATTTTGTGCTGCCAGATGATGATCAAGGAT 840

```

```

841 TGTGTGGCAAAAGAAATTTCTGTATCAGACAGATTTCTCAGACAACTTGTATTATACC 900
841 TGTGTGGCAAAAGAAATTTCTGTATCAGACAGATTTCTCAGACAACTTGTATTATACC 900

901 TTTGTGGCCAGGGTGAATTTCTGTATCAGATTTGTAGTAAATATCATCTTGGGTGGAT 960
901 TTTGTGGCCAGGGTGAATTTCTGTATCAGATTTGTAGTAAATATCATCTTGGGTGGAT 960

961 TCAACACCAACATATGAATCATGAGCGAGCAAGTGTATATGCTGCCACACTCTTGCTG 1020
961 TCAACACCAACATATGAATCATGAGCGAGCAAGTGTATATGCTGCCACACTCTTGCTG 1020

1021 GAAATCTGTGCAAAATATTTCTACATCTGAGAGCAGGAGTGAATTTCTGTTGAACTCCGG 1080
1021 GAAATCTGTGCAAAATATTTCTACATCTGAGAGCAGGAGTGAATTTCTGTTGAACTCCGG 1080

1081 CATTTGACTGGGAGTGAGACCAAAATCTGGAATAATGCAATCAGCAAACTCTCTGTAA 1140
1081 CATTTGACTGGGAGTGAGACCAAAATCTGGAATAATGCAATCAGCAAACTCTCTGTAA 1140

1141 GGTAAGAGTCAGAGATATGACCGTCCCTACAGCAATGTGGACAGGAGTCAAGACTGGC 1200
1141 GGTAAGAGTCAGAGATATGACCGTCCCTACAGCAATGTGGACAGGAGTCAAGACTGGC 1200

1201 TTTCAATCCAGAGAGCGTGAATATGCTCTCTGAGGTGACCACTCATCTACCATTA 1260
1201 TTTCAATCCAGAGAGCGTGAATATGCTCTCTGAGGTGACCACTCATCTACCATTA 1260

1261 AGAATATTTCTGAAATGGGCTCAGCTGGAATTTCTCTGGGTTTGGATGCTCTCACCGTA 1320
1261 AGAATATTTCTGAAATGGGCTCAGCTGGAATTTCTCTGGGTTTGGATGCTCTCACCGTA 1320

1321 TGTAAATGAATCATCTGATCTGATGAGGAGGAGCAACCTTTCCAGGGAGGCT 1380
1321 TGTAAATGAATCATCTGATCTGATGAGGAGGAGCAACCTTTCCAGGGAGGCT 1380

1381 GTGAGGCGGTATTTGTGAAGCATCTGACACTGACGATCTTAGGACCACTCTCTGAGGATG 1440
1381 GTGAGGCGGTATTTGTGAAGCATCTGACACTGACGATCTTAGGACCACTCTCTGAGGATG 1440

1441 GGGTAGAGCCCATGAGGAGCAAGATTTAGGAGAGCAGACCTTAGTATATCATTTTTCAGA 1500
1441 GGGTAGAGCCCATGAGGAGCAAGATTTAGGAGAGCAGACCTTAGTATATCATTTTTCAGA 1500

1501 TTCCCTGCACTTGGCACTTAATCCGACACTTTACATTTTCTGTAATTTAAA 1560
1501 TTCCCTGCACTTGGCACTTAATCCGACACTTTACATTTTCTGTAATTTAAA 1560

1561 GTACTTATAGGTAATAGAGGTTTGTATGCTATATATATATCTACCATCTTGAAGGGT 1620
1561 GTACTTATAGGTAATAGAGGTTTGTATGCTATATATATATCTACCATCTTGAAGGGT 1620

1621 AGGTTTACCTGATAGGCAAAATATCTAGACATCTCTATATCTATCTCAGGTAATCTC 1680
1621 AGGTTTACCTGATAGGCAAAATATCTAGACATCTCTATATCTATCTCAGGTAATCTC 1680

1681 TTTAAACACCTATTTGTTTTCTATAAGCCATATTTTGGAGCACTAAAGTAAATGGC 1740
1681 TTTAAACACCTATTTGTTTTCTATAAGCCATATTTTGGAGCACTAAAGTAAATGGC 1740

1741 AAATTTGGAGCAGATATTCAGGCTGAGGATCTGTGATTTATGTTGATTTTGACAAATAA 1800
1741 AAATTTGGAGCAGATATTCAGGCTGAGGATCTGTGATTTATGTTGATTTTGACAAATAA 1800

1801 GCTAGACATTTTCCCTTTGTTCCACAGAGACATAACACTACCTCAGGAAAGCTGAGCTGC 1860
1801 GCTAGACATTTTCCCTTTGTTCCACAGAGACATAACACTACCTCAGGAAAGCTGAGCTGC 1860

1861 TTTAAGGACAAACAAACAAATCAGTGTGTACAGTATGGAATCTATGTTAAGCATTT 1920
1861 TTTAAGGACAAACAAACAAATCAGTGTGTACAGTATGGAATCTATGTTAAGCATTT 1920

1921 CTCAGAAATAGGCCAAAGTTTATATAGTGTCTCTCAGGGAAGAAATTTTATAGGATGTTT 1980

```

Db 1921 CTCAGAAAGGCCAGTTTATAGTTCATCTCAGGAGAGAAATTTTATAGATGTTT 1980
Qy 1981 ATGAGTCTCCATAAATGCAATTCGCAATACATAAAAAAAGGCGG 2040
Db 1981 ATGAGTCTCCATAAATGCAATTCGCAATACATAAAAAAAGGCGG 2040
Qy 2041 CGCG 2044
Db 2041 CGCG 2044

RESULT 2
US-10-042-431-45
; Sequence 45, Application US/10042431
; Publication No. US20020182675A1
; GENERAL INFORMATION:
; APPLICANT: MCCARTHY, Sean A
; APPLICANT: BARNES, Thomas M
; APPLICANT: FRASER, Christopher C
; APPLICANT: SHARP, John D
; TITLE OF INVENTION: NOVEL GENES ENCODING PROTEINS HAVING DIAGNOSTIC,
; TITLE OF INVENTION: PREVENTIVE, THERAPEUTIC, AND OTHER USES
; FILE REFERENCE: 10147-6U2
; CURRENT APPLICATION NUMBER: US/10/042,431
; CURRENT FILING DATE: 2001-10-25
; PRIOR APPLICATION NUMBER: US 09/333,159
; PRIOR FILING DATE: 1999-06-14
; PRIOR APPLICATION NUMBER: US 09/578,063
; PRIOR FILING DATE: 2000-05-24
; NUMBER OF SEQ ID NOS: 79
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 45
; LENGTH: 2044
; TYPE: DNA
; ORGANISM: Homo sapiens
US-10-042-431-45

Query Match 100.0%; Score 2044; DB 14; Length 2044;
Best Local Similarity 100.0%; Pred. No. 0;
Matches 2044; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 GTCGACCCACGCGTCCGGGGAATTCAGCAGAGAAATATGTGAAGAGTTTTTAAACCCAC 60
Db 1 GTCGACCCACGCGTCCGGGGAATTCAGCAGAGAAATATGTGAAGAGTTTTTAAACCCAC 60

Qy 61 AAATCTTCTTACTTTAGAAATAGTTGTTAATTCGCGAGGAAAAAATAATGCAAGATGTT 120
Db 61 AAATCTTCTTACTTTAGAAATAGTTGTTAATTCGCGAGGAAAAAATAATGCAAGATGTT 120

Qy 121 GGACCATGTTGAAACCTTGTCAAGACAGTGGATTGTCTCACAGAAATGGAATGTGGC 180
Db 121 GGACCATGTTGAAACCTTGTCAAGACAGTGGATTGTCTCACAGAAATGGAATGTGGC 180

Qy 181 TTCTGATTTCTGTTGGCGTATATGTTCCAGAGAAATGTGAATTCAGTACATATGCCAACTA 240
Db 181 TTCTGATTTCTGTTGGCGTATATGTTCCAGAGAAATGTGAATTCAGTACATATGCCAACTA 240

Qy 241 AAGCTGTGGACCCAGCAAGCATTCATGAATATTAGTGAATCATCCAAATCAAGGCTATC 300
Db 241 AAGCTGTGGACCCAGAGCAATTCATGAATATTAGTGAATCATCCAAATCAAGGCTATC 300

Qy 301 CCTGTGAGGAATATCAAGTTCGCAACTGAAGATGGGTATATCCTTTCTGTTAAACAGGATTC 360
Db 301 CCTGTGAGGAATATCAAGTTCGCAACTGAAGATGGGTATATCCTTTCTGTTAAACAGGATTC 360

Qy 361 CTCGAGGCTAGTGCAACTAAGAGACAGGTTCCAGGCGCTGTGTTACTGACGATG 420
Db 361 CTCGAGGCTAGTGCAACTAAGAGACAGGTTCCAGGCGCTGTGTTACTGACGATG 420

Qy 421 GCCTAGTTGGAGGCTAGCAACTGGATTTCCAACTGCCCAACAATAGCCTGGGCTTCA 480
Db 421 GCCTAGTTGGAGGCTAGCAACTGGATTTCCAACTGCCCAACAATAGCCTGGGCTTCA 480

Qy 481 TTCTGCGACATGCTGGTTTTGACGTGTGATGGGAAACAGCAGGGGAAACCGCTGCTC 540
Db 481 TTCTGCGACATGCTGGTTTTGACGTGTGATGGGAAACAGCAGGGGAAACCGCTGCTC 540

Qy 541 GAAAAACAAGACACTCTCCATAGACCAAGATGAGTTCTGGGCTTTCAAGTTATGATGAGA 600
Db 541 GAAAAACAAGACACTCTCCATAGACCAAGATGAGTTCTGGGCTTTCAAGTTATGATGAGA 600

Qy 601 TGGCTAGGTTTGACCTTCTGCGATGATAAACTTTATTTGCGAGAAACGGGCCAGGAAA 660
Db 601 TGGCTAGGTTTGACCTTCTGCGATGATAAACTTTATTTGCGAGAAACGGGCCAGGAAA 660

Qy 661 AGATCTATTATGTCGCGCTATTCAAGGCGCACCACTGGGCTTTATTTGCAATTTTCCACCA 720
Db 661 AGATCTATTATGTCGCGCTATTCAAGGCGCACCACTGGGCTTTATTTGCAATTTTCCACCA 720

Qy 721 TGGCAGAGCTGGCTCAGAAAAATCAAAATGTAATTTGCTTTCAGCACCCTAGCCTGTGA 780
Db 721 TGGCAGAGCTGGCTCAGAAAAATCAAAATGTAATTTGCTTTCAGCACCCTAGCCTGTGA 780

Qy 781 AGCATGCAAAAAAGCCCGGACCAAAATTTGCTGTCGACAGATATGATGATCAAGGAT 840
Db 781 AGCATGCAAAAAAGCCCGGACCAAAATTTGCTGTCGACAGATATGATGATCAAGGAT 840

Qy 841 TGTGCGCAAAAAAGAAATTTCTGTATCAGACAGAAATTTCTCAGACAACTTGTGTTATACC 900
Db 841 TGTGCGCAAAAAAGAAATTTCTGTATCAGACAGAAATTTCTCAGACAACTTGTGTTATACC 900

Qy 901 TTTGTGGCCAGGTGATCTTGTATCAGATTTGATGATGATGATGATGATGATGATGATGAT 960
Db 901 TTTGTGGCCAGGTGATCTTGTATCAGATTTGATGATGATGATGATGATGATGATGATGAT 960

Qy 961 TCAACACCAAAATATGAACATGAGCCGAGCAAGTGTATATGCTGCCCACTCTGCTG 1020
Db 961 TCAACACCAAAATATGAACATGAGCCGAGCAAGTGTATATGCTGCCCACTCTGCTG 1020

Qy 1021 GAACATCTGTGCAAAATATTTCTACTGGAGCCGAGCAGTGAATTTCTGGTGAATCTCCGG 1080
Db 1021 GAACATCTGTGCAAAATATTTCTACTGGAGCCGAGCAGTGAATTTCTGGTGAATCTCCGG 1080

Qy 1081 CATTTGACTGGGGGAGTGAGACCAAAATCTCGAAAAATGCAATCAGCCAACTCTCTGTAA 1140
Db 1081 CATTTGACTGGGGGAGTGAGACCAAAATCTCGAAAAATGCAATCAGCCAACTCTCTGTAA 1140

Qy 1141 GGTACAGAGTCAAGATATGACGGTCCCTACAGCAATGTGACAGAGGTGAGGACTGGC 1200
Db 1141 GGTACAGAGTCAAGATATGACGGTCCCTACAGCAATGTGACAGAGGTGAGGACTGGC 1200

Qy 1201 TTTCAAAATCCAGAAAGCTGAAATGCTGCTCTCTGAGGTGACCAACCTCTATCTACCAT 1260
Db 1201 TTTCAAAATCCAGAAAGCTGAAATGCTGCTCTCTGAGGTGACCAACCTCTATCTACCAT 1260

Qy 1261 AGAATATCTGAAATGGGCTCAAGTGGATTTCAATCTGGGTTTGGATGCTCTCAGCGTA 1320
Db 1261 AGAATATCTGAAATGGGCTCAAGTGGATTTCAATCTGGGTTTGGATGCTCTCAGCGTA 1320

Qy 1321 TGTACAAATCAAAATCAATCTGATGACAGAGAGAGACCAACCTTTCCAGGAGCGT 1380
Db 1321 TGTACAAATCAAAATCAATCTGATGACAGAGAGAGACCAACCTTTCCAGGAGCGT 1380

Qy 1381 GTGAGGCCCTATGTAAGCATCTGACATGACAGTGTAGGACAACTCTCTGAGGAGTG 1440
Db 1381 GTGAGGCCCTATGTAAGCATCTGACATGACAGTGTAGGACAACTCTCTGAGGAGTG 1440

Qy 1441 GGGCTAGGACCCATGAAGGCAAAATAGCGAGAGAGAGACCTAGTATACATTTTTCAGA 1500
Db 1441 GGGCTAGGACCCATGAAGGCAAAATAGCGAGAGAGAGACCTAGTATACATTTTTCAGA 1500

Qy 1501 TTCCCTGCACTTGGCACTAAATCCGACACTTACATTTTACATTTTCTGTAATTTAA 1560
Db 1501 TTCCCTGCACTTGGCACTAAATCCGACACTTACATTTTACATTTTCTGTAATTTAA 1560

```

QY 1561 GTACTTATTAGGTAAATAGAGGTTTGTATGCTATTATATATATCTTACCATCTTGAAGGT 1620
Db 1561 GTACTTATTAGGTAAATAGAGGTTTGTATGCTATTATATATATCTTACCATCTTGAAGGT 1620
QY 1621 AGTTTTTACCTGATAGCCAGAAATATCTAGACATCTCTATATATCTTACCATCTTGAAGGT 1680
Db 1621 AGTTTTTACCTGATAGCCAGAAATATCTAGACATCTCTATATATCTTACCATCTTGAAGGT 1680
QY 1681 TTTAAACACCTATCTTCTTCTATAGCCATATTTTGGAGCCTAAAGTAAATGCG 1740
Db 1681 TTTAAACACCTATCTTCTTCTATAGCCATATTTTGGAGCCTAAAGTAAATGCG 1740
QY 1741 AAATTGGGACAGATATTGAGGTCTGGAGTCTGGAGTATTTGTGATTTTGAACAAATAA 1800
Db 1741 AAATTGGGACAGATATTGAGGTCTGGAGTCTGGAGTATTTGTGATTTTGAACAAATAA 1800
QY 1801 GCTAGACATTTTACCTTGTGTCACAGAGACATACATCTCTAGGAGCTGAGCTGC 1860
Db 1801 GCTAGACATTTTACCTTGTGTCACAGAGACATACATCTCTAGGAGCTGAGCTGC 1860
QY 1861 TTTAAGGACAAACAAACAAATCAGTGTTCACAGTATGATGAAATCTATGTTAAGCAAT 1920
Db 1861 TTTAAGGACAAACAAACAAATCAGTGTTCACAGTATGATGAAATCTATGTTAAGCAAT 1920
QY 1921 CTGAGATAGGCAAGTTTATAGTGTGATCTCTAGGAGGAGAAATTTTATAGGATGTT 1980
Db 1921 CTGAGATAGGCAAGTTTATAGTGTGATCTCTAGGAGGAGAAATTTTATAGGATGTT 1980
QY 1981 ATGAGTCTTCCATTAATGATCTCTGATTTACATATACATATATATATATATATATAT 2040
Db 1981 ATGAGTCTTCCATTAATGATCTCTGATTTACATATACATATATATATATATATATAT 2040
QY 2041 CCGC 2044
Db 2041 CCGC 2044

```

RESULT 3

```

US-09-835-996A-20
; Sequence 20, Application US/09835996A
; Patent No. US20020142953A1
; GENERAL INFORMATION:
; APPLICANT: Ballinger, Dennis
; APPLICANT: Loeb, Debra
; APPLICANT: Montgomery, Julie
; APPLICANT: Tang, Y. Tom
; APPLICANT: Zhou, Ping
; APPLICANT: Goodrich, Ryle
; APPLICANT: Liu, Chenghua
; APPLICANT: Asundi, Vinod
; APPLICANT: Zhao, Qing
; APPLICANT: Wehrman, Tom
; APPLICANT: Dumanac, Radoje
; APPLICANT: Ren, Feiyan
; APPLICANT: Qian, Xiaohong
; APPLICANT: Wang, Dunru
; TITLE OF INVENTION: MATERIALS AND METHODS RELATING TO LIPID METABOLISM
; FILE REFERENCE: 28110/35915A
; CURRENT APPLICATION NUMBER: US/09/835,996A
; CURRENT FILING DATE: 2001-04-16
; PRIOR FILING DATE: 2000-04-14
; PRIOR FILING DATE: 2000-04-14
; PRIOR FILING DATE: 2000-11-17
; PRIOR FILING DATE: 2000-09-22
; PRIOR FILING DATE: 2000-08-03
; PRIOR FILING DATE: 2000-06-20
; NUMBER OF SEQ ID NOS: 45
; SOFTWARE: PatentIn version 3.0
; SEQ ID NO 20

```

```

; LENGTH: 2035
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: CDS
; LOCATION: (189)..(1415)
US-09-835-996A-20

Query Match
Best Local Similarity 97.5%; Score 1993.8; DB 10; Length 2035;
Matches 1995; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 18 GGGAAATTCAGCAGGAAATATCTGAAGAGTATTTTAAACCCACAAATCTTCTTACTTTA 77
Db 39 GGGAAATTCAGCAGGAAATATCTGAAGAGTATTTTAAACCCACAAATCTTCTTACTTTA 98
QY 78 GAATTTAGTCTTCAATTTGGCAGGAAATATTAATGAGATTTTGGACCATGTTGGAACACC 137
Db 99 GAATTTAGTCTTCAATTTGGCAGGAAATATTAATGAGATTTTGGACCATGTTGGAACACC 158
QY 138 TTGTCAAGACAGTGGATTTGTTCCACACAGAAATGGAATTTGGCTTCTGATTTCTGTGGCG 197
Db 159 TTGTCAAGACAGTGGATTTGTTCCACACAGAAATGGAATTTGGCTTCTGATTTCTGTGGCG 218
QY 198 TATATGTTCCAGAGAAATGTAATTCAGTACATATGCAACTAAAGCTCTGACCCAGAA 257
Db 219 TATATGTTCCAGAGAAATGTAATTCAGTACATATGCAACTAAAGCTCTGACCCAGAA 278
QY 258 GCATTTCATGAATATTTAGTGAATCATCCAAATCAAGAGCTATCCCTGTGAGGAATATGAA 317
Db 279 GCATTTCATGAATATTTAGTGAATCATCCAAATCAAGAGCTATCCCTGTGAGGAATATGAA 338
QY 318 GTCCAACTGAAGATGGGTATATCTTCTGTTTAAAGAGATTCCTCGAGGCTTAGTGCAA 377
Db 339 GTCCAACTGAAGATGGGTATATCTTCTGTTTAAAGAGATTCCTCGAGGCTTAGTGCAA 398
QY 378 CCTAAGAACAGAGTTCAGGCTCTGGTGTGTTACTGAGCATGGCTTAGTGTGAGGTGCT 437
Db 399 CCTAAGAACAGAGTTCAGGCTCTGGTGTGTTACTGAGCATGGCTTAGTGTGAGGTGCT 458
QY 438 AGCAATCTGATTTCCAACTGCTGCAACATAGCTGGCTTCACTTCTGAGATGCTGCT 497
Db 459 AGCAATCTGATTTCCAACTGCTGCAACATAGCTGGCTTCACTTCTGAGATGCTGCT 518
QY 498 TTTGACCTGTGATGGGAAACAGCAGGGGAAACCGCTGCTCGAAACACACAGACACTC 557
Db 519 TTTGACCTGTGATGGGAAACAGCAGGGGAAACCGCTGCTCGAAACACACAGACACTC 578
QY 558 TCCATAGACCAAGATGATGCTGGGCTTTCAGTTATGATGAGATGGCTTAGTGTGACCTT 617
Db 579 TCCATAGACCAAGATGATGCTGGGCTTTCAGTTATGATGAGATGGCTTAGTGTGACCTT 638
QY 618 CCTGACGTGATTAACCTTTATTTTGCAGAAACCGGCGCAGGAAAGATCTATTATGTGCGC 677
Db 639 CCTGACGTGATTAACCTTTATTTTGCAGAAACCGGCGCAGGAAAGATCTATTATGTGCGC 698
QY 678 TATTACAGGCGCACCATGCTGCTTTATTCGATTTTCCACCATGCGAGAGTGGCTCAG 737
Db 699 TATTACAGGCGCACCATGCTGCTTTATTCGATTTTCCACCATGCGAGAGTGGCTCAG 758
QY 738 AAATTCAAATGATTTTGTCTTTAGCCACCATGCTGTTAAGCATGCAAAAGCCCC 797
Db 759 AAATTCAAATGATTTTGTCTTTAGCCACCATGCTGTTAAGCATGCAAAAGCCCC 818
QY 798 GGGACCAAAATTTTGTGCTGCCAGATATGATGATCAAGGGATTTGTTGCAAAAGAA 857
Db 819 GGGACCAAAATTTTGTGCTGCCAGATATGATGATCAAGGGATTTGTTGCAAAAGAA 878
QY 858 TTTCTGTATCAGACCAATTTCTCAGACAACTTTGTTATTTTACCTTTGTGCGCAGGTGATT 917
Db 879 TTTCTGTATCAGACCAATTTCTCAGACAACTTTGTTATTTTACCTTTGTGCGCAGGTGATT 938
QY 918 CTTCATCAGATTTTGTAGTAATATCATGTTACTTCTGGGTGGATTCACACCAATATG 977

```

Db 939 CTTGATCAGATTGTTAGTAATATCATGTTACTTCTGGTGGATTCAACACCAACATATG 998
 Qy 978 AACATGAGCGAGCAAGTGTATATGCTGCCCCACACTCTTGTGGAACATCTGCAAAAT 1037
 Db 999 AACATGAGCGAGCAAGTGTATATGCTGCCCCACACTCTTGTGGAACATCTGCAAAAT 1058
 Qy 1038 ATTCTACATGAGCGAGCAAGTGTATATGCTGCCCCACACTCTTGTGGAACATCTGCAAAAT 1097
 Db 1059 ATTCTACATGAGCGAGCAAGTGTATATGCTGCCCCACACTCTTGTGGAACATCTGCAAAAT 1118
 Qy 1098 GAGACCAAAATCTGGAABATGCAATGCAACGCAACCTCTGTAAGGTACAGAGTCAAGAT 1157
 Db 1119 GAGACCAAAATCTGGAABATGCAATGCAACGCAACCTCTGTAAGGTACAGAGTCAAGAT 1178
 Qy 1158 ATGACGCTCCCTACAGCAATGTCAGCAGGAGGTGAGGAGTGGCTTTCAAAATCAGAGAC 1217
 Db 1179 ATGACGCTCCCTACAGCAATGTCAGCAGGAGGTGAGGAGTGGCTTTCAAAATCAGAGAC 1238
 Qy 1218 GTGAAATGCTGCTCTGAGGTGACCAACCTCTGTAAGGTACAGAGTCAAGAT 1277
 Db 1239 GTGAAATGCTGCTCTGAGGTGACCAACCTCTGTAAGGTACAGAGTCAAGAT 1298
 Qy 1278 GCTCAGCTGGATTTCACTCGGGTTTGGATGCTCCTCACCGTATGACATGAAATCATC 1337
 Db 1299 GCTCAGCTGGATTTCACTCGGGTTTGGATGCTCCTCACCGTATGACATGAAATCATC 1358
 Qy 1338 CATCTGATGAGCGAGGAGACCAACCTTCCAGGAGCGGTGAGGCGGTATTGTA 1397
 Db 1359 CATCTGATGAGCGAGGAGACCAACCTTCCAGGAGCGGTGAGGCGGTATTGTA 1418
 Qy 1398 AGCATCTGACATGACGATCTTAGGACCAACCTCTGAGGAGTGGGTAGGACCCATGAA 1457
 Db 1419 AGCATCTGACATGACGATCTTAGGACCAACCTCTGAGGAGTGGGTAGGACCCATGAA 1478
 Qy 1458 GCGAATATGAGGAGGAGACGATGATATATATTTTCCAGATTCCTGCACTTGGCAC 1517
 Db 1479 GCGAATATGAGGAGGAGACGATGATATATATTTTCCAGATTCCTGCACTTGGCAC 1538
 Qy 1518 TAAATCCGACATTTACATTTTACATTTTCTGTAATTTAAAGTACTTATTAGGTAAT 1577
 Db 1539 TAAATCCGACATTTACATTTTCTGTAATTTAAAGTACTTATTAGGTAAT 1598
 Qy 1578 AGAGTTTGTATGCTATATATATTTTACATCTTGAAGGTTAGTTTACCTGATAGC 1637
 Db 1599 AGAGTTTGTATGCTATATATATTTTACATCTTGAAGGTTAGTTTACCTGATAGC 1658
 Qy 1638 CAGAAATATCTAGACATCTCTATATATCTCAGGTAATCTCTTTAAACACCTATTCT 1697
 Db 1659 CAGAAATATCTAGACATCTCTATATATCTCAGGTAATCTCTTTAAACACCTATTCT 1718
 Qy 1698 TTTTCTATAGCCATATTTTGGAGCATTAAGTAAATGGAATTTGGACAGATAT 1757
 Db 1719 TTTTCTATAGCCATATTTTGGAGCATTAAGTAAATGGAATTTGGACAGATAT 1778
 Qy 1758 GAGTCTGGAGTCTGTGGATATTTGTTGACTTTTGAACAAATTAAGTAAATTTTCACT 1817
 Db 1779 GAGTCTGGAGTCTGTGGATATTTGTTGACTTTTGAACAAATTAAGTAAATTTTCACT 1838
 Qy 1818 TGTGTCACAGACATAACTACCTCAGGAGGCTGAGTCTTTTAAAGGACAAACAA 1877
 Db 1839 TGTGTCACAGACATAACTACCTCAGGAGGCTGAGTCTTTTAAAGGACAAACAA 1898
 Qy 1878 CAAATCAGTGTATGAGTGGATGAAATCTATGTTAAAGATCTCAGATTAAGGCCAAG 1937
 Db 1899 CAAATCAGTGTATGAGTGGATGAAATCTATGTTAAAGATCTCAGATTAAGGCCAAG 1958
 Qy 1938 TTTTATAGTTGCACTCAGGAGGAAATTTTATAGGATGTTTATGAGTCTTCAATAAA 1997
 Db 1959 TTTTATAGTTGCACTCAGGAGGAAATTTTATAGGATGTTTATGAGTCTTCAATAAA 2018
 Qy 1998 TGAATCTGCAATACAT 2014

Db 2019 TGCATTCTGCATTACAT 2035
 RESULT 4
 ; Sequence 9, Application US/09835996A
 ; Patent No. US20020142953A1
 ; GENERAL INFORMATION:
 ; APPLICANT: Ballinger, Dennis
 ; APPLICANT: Loeb, Debra
 ; APPLICANT: Montgomery, Julie
 ; APPLICANT: Tang, Y. Tom
 ; APPLICANT: Zhou, Ping
 ; APPLICANT: Goodrich, Ryle
 ; APPLICANT: Liu, Chenchua
 ; APPLICANT: Asundi, Vinod
 ; APPLICANT: Zhao, Qing
 ; APPLICANT: Wehrman, Tom
 ; APPLICANT: Drmanac, Radoje
 ; APPLICANT: Ren, Feiyan
 ; APPLICANT: Qian, Xiahong
 ; APPLICANT: Wang, Dunrui
 ; TITLE OF INVENTION: MATERIALS AND METHODS RELATING TO LIPID METABOLISM
 ; FILE REFERENCE: 28110/35915A
 ; CURRENT APPLICATION NUMBER: US/09/835,996A
 ; CURRENT FILING DATE: 2001-04-16
 ; PRIOR APPLICATION NUMBER: US 60/197,137
 ; PRIOR FILING DATE: 2000-04-14
 ; PRIOR APPLICATION NUMBER: US 09/714,936
 ; PRIOR FILING DATE: 2000-11-17
 ; PRIOR APPLICATION NUMBER: US 09/667,298
 ; PRIOR FILING DATE: 2000-09-22
 ; PRIOR APPLICATION NUMBER: US 09/631,451
 ; PRIOR FILING DATE: 2000-08-03
 ; PRIOR APPLICATION NUMBER: US 09/598,042
 ; PRIOR FILING DATE: 2000-06-20
 ; NUMBER OF SEQ ID NOS: 45
 ; SOFTWARE: PatentIn version 3.0
 ; SEQ ID NO 9
 ; LENGTH: 1384
 ; TYPE: DNA
 ; ORGANISM: Homo sapiens
 ; US-09-835-996A-9
 Query Match 65.7%; Score 1343.4; DB 10; Length 1384;
 Best Local Similarity 99.9%; Pred. No. 0;
 Matches 1344; Conservative 0; Mismatches 1; Indels 0; Gaps 0;
 Qy 18 GGGAAATTCAGCAGGAAAATATGTGAAGAGTTTAAACCCACAAATTTCTTCTACTTTA 77
 Db 39 GGGAAATTCAGCAGGAAAATATGTGAAGAGTTTAAACCCACAAATTTCTTCTACTTTA 98
 Qy 78 GAAATTTAGTTTACATTTGGCAGGAAAATATAATGACAGATGTTGGACCATGTTGGAAC 137
 Db 99 GAAATTTAGTTTACATTTGGCAGGAAAATATAATGACAGATGTTGGACCATGTTGGAAC 158
 Qy 138 TTGTCAGACAGTGGATGCTCTACACAGAAATGGAATTTGCTTCTGATTTCTGGTGGCG 197
 Db 159 TTGTCAGACAGTGGATGCTCTACACAGAAATGGAATTTGCTTCTGATTTCTGGTGGCG 218
 Qy 198 TATATGTTCCAGAGAAATGGAATTCAGTACATATGCAACTAAAGCTGTGGACCCAGAA 257
 Db 219 TATATGTTCCAGAGAAATGGAATTCAGTACATATGCAACTAAAGCTGTGGACCCAGAA 278
 Qy 258 GCATTCATGAAATATAGTGAATCATCAACATCAAGGCTATCCCTGTGAGGAATATGAA 317
 Db 279 GCATTCATGAAATATAGTGAATCATCAACATCAAGGCTATCCCTGTGAGGAATATGAA 338
 Qy 318 GTCCCACTGAGATGGTATATCTTCTGTTTAAAGGATTTCTCGAGGCTAGTGCAA 377
 Db 339 GTCCCACTGAGATGGTATATCTTCTGTTTAAAGGATTTCTCGAGGCTAGTGCAA 398
 Qy 378 CCTAAGACAGAGTTCAGGCGCTGTGGTGTACTGACAGCATGGCCTTAGTTGGAGGTGCT 437

Db CCTAAGAGAGAGAGGTTCCAGGCTGTGGTGTATTCTGAGCATGCGCTAGTTGGAGTGCT 458
 QY AGCAACTGAGATTCCAACTGCGCCCAACATAGAGCTGGCTTCATCTCGGAGATGCTGGT 497
 Db AGCAACTGAGATTCCAACTGCGCCCAACATAGAGCTGGCTTCATCTCGGAGATGCTGGT 518
 QY TTTGACGTGTGGATGGGGAACAGACAGAGGGGAAACGCTGCTGCTCAAAACACAGACACTC 557
 Db TTTGACGTGTGGATGGGGAACAGACAGAGGGGAAACGCTGCTGCTCAAAACACAGACACTC 578
 QY TCCATAGACCAAGATGAGTCTGGCTTTTCAGTTATAGATGAGATGGCTAGCTTGGACCTT 617
 Db TCCATAGACCAAGATGAGTCTGGCTTTTCAGTTATAGATGAGATGGCTAGCTTGGACCTT 638
 QY CTGCAAGTGAATAACTTTTATTTTGCAGAAACCGGCGCAGGAAAGATCTATTATGTCGGC 677
 Db CTGCAAGTGAATAACTTTTATTTTGCAGAAACCGGCGCAGGAAAGATCTATTATGTCGGC 698
 QY TATTCACAGGGCACCAATGGGCTTTATGCAATTTTCCACCATGCGAGAGCTGGCTCAG 737
 Db TATTCACAGGGCACCAATGGGCTTTATGCAATTTTCCACCATGCGAGAGCTGGCTCAG 758
 QY AAAATCAAATGTATTTTGTCTTATGACACCCATAGCCTGTTAAGCATGCAAAAGAGCCCC 797
 Db AAAATCAAATGTATTTTGTCTTATGACACCCATAGCCTGTTAAGCATGCAAAAGAGCCCC 818
 QY GGGACCAAATTTTGTCTGCTCCAGATATGATGATCAAGGATTTTGGCAAAAGAA 857
 Db GGGACCAAATTTTGTCTGCTCCAGATATGATGATCAAGGATTTTGGCAAAAGAA 878
 QY TTTCTGTATCAGACAGATTTCTCAGACCACTGCTGTTATTTACCTTTTGGCCAGGTGATT 917
 Db TTTCTGTATCAGACAGATTTCTCAGACCACTGCTGTTATTTACCTTTTGGCCAGGTGATT 938
 QY CTGTATCAGATTGTTAGTAAATCATGTTCTCTGGGTGAGTCAACACCAACCAATATG 977
 Db CTGTATCAGATTGTTAGTAAATCATGTTCTCTGGGTGAGTCAACACCAACCAATATG 998
 QY AACATGAGCCGAGCAAGTGTATGCTGCCACACTCTGCTGGAACATCTGTGCAAAAT 1037
 Db AACATGAGCCGAGCAAGTGTATGCTGCCACACTCTGCTGGAACATCTGTGCAAAAT 1058
 QY ATTCTACACTGAGCCAGCAGTGAATTTCTGTTGAATCTCCGGGCAATTTGACTGGGGAGT 1097
 Db ATTCTACACTGAGCCAGCAGTGAATTTCTGTTGAATCTCCGGGCAATTTGACTGGGGAGT 1118
 QY GAGACCAAAATCTGGAAAAATGCAATCAGCCAACTCTGTAGAGTACAGATCAGAGAT 1157
 Db GAGACCAAAATCTGGAAAAATGCAATCAGCCAACTCTGTAGAGTACAGATCAGAGAT 1178
 QY ATGACGGTCCCTACAGCAATGTGGACAGAGGTGAGGAGGTGAGGAGTCCAGGAGT 1217
 Db ATGACGGTCCCTACAGCAATGTGGACAGAGGTGAGGAGGTGAGGAGTCCAGGAGT 1238
 QY GTGAAAAATCTGCTCTGAGGTGACCAACCTCATCTACCAATAGAAATTTCTGAAATGG 1277
 Db GTGAAAAATCTGCTCTGAGGTGACCAACCTCATCTACCAATAGAAATTTCTGAAATGG 1298
 QY GCTCAGCTGGAATTCATCTGGGGTTTGGATGCTCTCACCCTGATGTACAAATGAAATCATC 1337
 Db GCTCAGCTGGAATTCATCTGGGGTTTGGATGCTCTCACCCTGATGTACAAATGAAATCATC 1358
 QY CATCTGATCAGAGAGAGAGACCA 1362
 Db CATCTGATCAGAGAGAGAGACCA 1383

RESULT 5

US-09-759-1308-416
 ; Sequence 416, Application US/097591308
 ; Publication No. US2003002279A1
 ; GENERAL INFORMATION:

APPLICANT: Millennium Pharmaceuticals, Inc.
 APPLICANT: McCarthy, Sean A
 APPLICANT: Fraser, Christopher C
 APPLICANT: Sharp, John D
 APPLICANT: Barnes, Thomas S
 APPLICANT: Kirt, Susan J
 APPLICANT: Mackay, Charles R
 APPLICANT: Myers, Paul S
 APPLICANT: Leiby, Kevin R
 APPLICANT: Wrighton, Nicolas
 APPLICANT: Goodearl, Andrew
 APPLICANT: Holtzman, Douglas A
 TITLE OF INVENTION: NOVEL GENES ENCODING PROTEINS HAVING
 TITLE OF INVENTION: PROGNOSTIC, DIAGNOSTIC, PREVENTIVE, THERAPEUTIC, AND OTHER
 TITLE OF INVENTION: USES
 FILE REFERENCE: MPI00-5350NMIM
 CURRENT APPLICATION NUMBER: US/09/759,130B
 PRIOR FILING DATE: 2002-09-16
 PRIOR APPLICATION NUMBER: US 09/479,249
 PRIOR FILING DATE: 2000-01-07
 PRIOR APPLICATION NUMBER: US 09/559,497
 PRIOR FILING DATE: 2000-04-27
 PRIOR APPLICATION NUMBER: US 09/578,063
 PRIOR FILING DATE: 2000-05-24
 PRIOR APPLICATION NUMBER: US 09/333,159
 PRIOR FILING DATE: 1999-06-14
 PRIOR APPLICATION NUMBER: US 09/596,194
 PRIOR FILING DATE: 2000-07-14
 PRIOR APPLICATION NUMBER: US 09/342,364
 PRIOR FILING DATE: 1999-06-29
 PRIOR APPLICATION NUMBER: US 09/608,452
 PRIOR FILING DATE: 2000-06-30
 PRIOR APPLICATION NUMBER: US 09/393,996
 PRIOR FILING DATE: 1999-09-10
 PRIOR APPLICATION NUMBER: US 09/602,871
 PRIOR FILING DATE: 2000-06-23
 PRIOR APPLICATION NUMBER: US 09/420,707
 PRIOR FILING DATE: 1999-10-19
 NUMBER OF SEQ ID NOS: 460
 SOFTWARE: FASTSEQ for Windows Version 4.0
 SEQ ID NO 416
 LENGTH: 1269
 TYPE: DNA
 ORGANISM: Homo sapiens
 US-09-759-1308-416

Query Match 62.1%; Score 1269; DB 11; Length 1269;
 Best Local Similarity 100.0%; Pred. No. 0;
 Matches 1269; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 126 ATGTTGAAACCTTGTCAAGACAGTGGATTGTCTCAGCAGAAATGGAATGCGCTTCTG 185
 Db 1 ATGTTGAAACCTTGTCAAGACAGTGGATTGTCTCAGCAGAAATGGAATGCGCTTCTG 60

QY 186 ATCTGTCGCTATATGTTCCAGAGAAATGGAATTCAGTACATATGCCAATTAAGCT 245
 Db 61 ATCTGTCGCTATATGTTCCAGAGAAATGGAATTCAGTACATATGCCAATTAAGCT 120

QY 246 GTGACCCAGAGAGCTTCATGAATATTAGTGAATCATCAACATCAAGGCTATCCCTGT 305
 Db 121 GTGACCCAGAGAGCTTCATGAATATTAGTGAATCATCAACATCAAGGCTATCCCTGT 180

QY 306 GAGGAATATGAAGTCCAACTAAGAGACAGGTTCCAGGCTGTGGTGTACTGAGCATGGCCTA 365
 Db 181 GAGGAATATGAAGTCCAACTAAGAGACAGGTTCCAGGCTGTGGTGTACTGAGCATGGCCTA 240

QY 366 GGCCTAGTCAACCTAAGAGACAGGTTCCAGGCTGTGGTGTACTGAGCATGGCCTA 425
 Db 241 GGCCTAGTCAACCTAAGAGACAGGTTCCAGGCTGTGGTGTACTGAGCATGGCCTA 300

QY 426 GTTGGAGGTGCTAGCAACTGGAATTTCCAACTGCCCAACAATAGCCTGGGCTTTCATCTG 485
 Db 301 GTTGGAGGTGCTAGCAACTGGAATTTCCAACTGCCCAACAATAGCCTGGGCTTTCATCTG 360


```

Db      721  GGCACAAAAGAAATTTCTGTATCAGACAGATTTCTCAGACAACTTGTATTTACCTTTGT 780
Qy      906  GGCAGGTGATTTCTGTATCAGATTTGTAGTAATATCATGTTACTTCTCGGTGATTCATCAAC 965
Db      781  GGCAGGTGATTTCTGTATCAGATTTGTAGTAATATCATGTTACTTCTCGGTGATTCATCAAC 840
Qy      966  ACCAACATATGAACATGAGCCGACGCAAGTGTATATGCTGCCACACACTCTTGCTGGAACA 1025
Db      841  ACCAACATATGAACATGAGCCGACGCAAGTGTATATGCTGCCACACACTCTTGCTGGAACA 900
Qy      1026  TCTGTGCAAAATATTTACACTGAGCCGACGCAAGTGTATATGCTGCCACACTCTTGCTGGAACA 1085
Db      901  TCTGTGCAAAATATTTACACTGAGCCGACGCAAGTGTATATGCTGCCACACTCTTGCTGGAACA 960
Qy      1086  GACTGGGGGAGTGAGACCAAAATCTGCAAAATGCAATCAGCAAACTCTGTAAGGTAC 1145
Db      961  GACTGGGGGAGTGAGACCAAAATCTGCAAAATGCAATCAGCAAACTCTGTAAGGTAC 1020
Qy      1146  AGAGTCAGAGATATGAGCGTCCCTACAGCAATGTGGACGAGGTGAGACTGGCTTTCA 1205
Db      1021  AGAGTCAGAGATATGAGCGTCCCTACAGCAATGTGGACGAGGTGAGACTGGCTTTCA 1080
Qy      1206  AATCCAGAGAGCGTGAATATGCTCTCTGAGGTGACCAACCTCATCTACCAATAGAAAT 1265
Db      1081  AATCCAGAGAGCGTGAATATGCTCTCTGAGGTGACCAACCTCATCTACCAATAGAAAT 1140
Qy      1266  ATTCTCTGAATGGGCTCAGTGGAATTTCTCTGAGGTGATGCTCTCCTCAGCGATGTATAC 1325
Db      1141  ATTCTCTGAATGGGCTCAGTGGAATTTCTCTGAGGTGATGCTCTCCTCAGCGATGTATAC 1200
Qy      1326  AATGAATATCATCTGATGAGCAGGAGGAGACCAACCTTTCCAGGAGCGGTGTAG 1385
Db      1201  AATGAATATCATCTGATGAGCAGGAGGAGACCAACCTTTCCAGGAGCGGTGTAG 1260
Qy      1386  GCCGTATTTG 1394
Db      1261  GCCGTATTTG 1269

```

RESULT 7

```

US-09-811-825-1
; Sequence 1, Application US/09811825
; Patent No. US20020144297A1
; GENERAL INFORMATION:
; APPLICANT: YAN, Chunhua et al.
; TITLE OF INVENTION: ISOLATED HUMAN LIPASE PROTEINS, NUCLEIC
; MOLECULES ENCODING HUMAN LIPASE PROTEINS, AND USES
; THEREOF
; FILE REFERENCE: CLO01170
; CURRENT APPLICATION NUMBER: US/09/811.825
; CURRENT FILING DATE: 2001-03-20
; NUMBER OF SEQ ID NOS: 4
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 1
; LENGTH: 1206
; TYPE: DNA
; ORGANISM: Human
US-09-811-825-1

```

```

Query Match      55.1%; Score 1126; DB 10; Length 1206;
Best Local Similarity 100.0%; Pred. No. 3e-298;
Matches 1126; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
Qy      273  AGTGAATATCCAACTCAAGGTATCCCTGTGAGGAATATGAATCGCAACTGAAGAT 332
Db      76  AGTGAATATCCAACTCAAGGTATCCCTGTGAGGAATATGAATCGCAACTGAAGAT 135
Qy      333  GGGTATATCCCTTTCTGTATTAACAGGATTCCTGAGGCTAGTGCAACCTAAGAGACAGGT 392
Db      136  GGGTATATCCCTTTCTGTATTAACAGGATTCCTGAGGCTAGTGCAACCTAAGAGACAGGT 195
Qy      393  TCCAGGCTGTGGTGTACTGCAGATGGCCCTAGTTGAGGTGCTAGCAACTGATTTCC 452

```

RESULT 8

```

US-09-811-825-3
; Sequence 3, Application US/09811825
; Patent No. US20020144297A1
; GENERAL INFORMATION:
; APPLICANT: YAN, Chunhua et al.

```

```

Db      196  TCCAGGCTGTGGTGTATTCTGACATGCGCTAGTTGGAGGTGCTAGCAACTGGATTTCC 255
Qy      453  AACCTGCCCAACAATAGCTGGGCTTCATTTCTGGCAGATGCTGGTGTGACGTGTGGATG 512
Db      256  AACCTGCCCAACAATAGCTGGGCTTCATTTCTGGCAGATGCTGGTGTGACGTGTGGATG 315
Qy      513  GGGAAACACGAGGGGAAACGGCTCTCGAATAACACACAGACACTCTCCATAGACCAAGAT 572
Db      316  GGGAAACACGAGGGGAAACGGCTCTCGAATAACACACAGACACTCTCCATAGACCAAGAT 375
Qy      573  GAGTCTTGGGCTTTTCACTATGATGATGCTAGGTGTTGACCTTCTCGCAGTGTATAAC 632
Db      376  GAGTCTTGGGCTTTTCACTATGATGATGCTAGGTGTTGACCTTCTCGCAGTGTATAAC 435
Qy      633  TTTATTTTGCAGAAAAACGGGCGCAGGAAAAAGATCTATTATGCTCGCTATTTCACAGGGCACC 692
Db      436  TTTATTTTGCAGAAAAACGGGCGCAGGAAAAAGATCTATTATGCTCGCTATTTCACAGGGCACC 495
Qy      693  ACCATGGGCTTTTATGATGATGCTAGGTGTTGACCTTCTCGCAGTGTGTATTAAC 752
Db      496  ACCATGGGCTTTTATGATGATGCTAGGTGTTGACCTTCTCGCAGTGTGTATTAAC 555
Qy      753  TTTGCTTTAGCACCATAGCCACTGTTAAGCATGCAAAAAAGCCCGGACCAAAATTTTG 812
Db      556  TTTGCTTTAGCACCATAGCCACTGTTAAGCATGCAAAAAAGCCCGGACCAAAATTTTG 615
Qy      813  TTGCTGCCAGATATGATGATGCTAGGTGTTGCAAAAAAGATTTCTGTATCAGACACC 872
Db      616  TTGCTGCCAGATATGATGATGCTAGGTGTTGCAAAAAAGATTTCTGTATCAGACACC 675
Qy      873  AGATTTCTCAGACCAACTGTTATTTAGCTTTGGCCAGGTGATTTCTGTATCAGATTTGT 932
Db      676  AGATTTCTCAGACCAACTGTTATTTAGCTTTGGCCAGGTGATTTCTGTATCAGATTTGT 735
Qy      933  AGTAAATATCATCTTCTGGGTGATTTCAACCAACAATATGAACATGAGCCGAGCA 992
Db      736  AGTAAATATCATCTTCTGGGTGATTTCAACCAACAATATGAACATGAGCCGAGCA 795
Qy      993  AGTGATATGCTGCCACACTCTTCTGGAACATCTCTGCAAAATATTTCTACACTGGAGC 1052
Db      796  AGTGATATGCTGCCACACTCTTCTGGAACATCTCTGCAAAATATTTCTACACTGGAGC 855
Qy      1053  CAGGAGTGAATTTCTGGTGAATCTCGGGCAATTTGACCTGGGGGAGTGAGACCAAAATCTG 1112
Db      856  CAGGAGTGAATTTCTGGTGAATCTCGGGCAATTTGACCTGGGGGAGTGAGACCAAAATCTG 915
Qy      1113  GAAAAATGCAATCAGCCAACTCTCTGTAAGGTACAGAGTACAGATATGACGGTCCCTACA 1172
Db      916  GAAAAATGCAATCAGCCAACTCTCTGTAAGGTACAGAGTACAGATATGACGGTCCCTACA 975
Qy      1173  GCAATGTGGAAGGAGGTGAGGACTGGCTTTCAATCCAGAGACGTGAAATCTGCTC 1232
Db      976  GCAATGTGGAAGGAGGTGAGGACTGGCTTTCAATCCAGAGACGTGAAATCTGCTC 1035
Qy      1233  TCTGAGGTGACCAACCTCATCTACCAATAAGATATTTCTGAATGGGCTCAGCTGATTTTC 1292
Db      1036  TCTGAGGTGACCAACCTCATCTACCAATAAGATATTTCTGAATGGGCTCAGCTGATTTTC 1095
Qy      1293  ATCTGGGCTTTGGATGCTCTCAGCGTATGTAACAATAATCAATCAATCAATCAATCAATCA 1352
Db      1096  ATCTGGGCTTTGGATGCTCTCAGCGTATGTAACAATAATCAATCAATCAATCAATCAATCA 1155
Qy      1353  GAGGAGACCAACCTTTCCAGGAGCGGTGTCAGGCGGTATTGTGAA 1398
Db      1156  GAGGAGACCAACCTTTCCAGGAGCGGTGTCAGGCGGTATTGTGAA 1201

```


Db	475	ACAGCTGTGGATCCTGAAA	CAAA	CATGAATGTGGTGAATATCTCTTACTCGGGATTC	534
Qy	300	CCCTGTGAGGAATATGAAGT	CGCAACTG	AAGATGGGTATATCCTTTCTGTAA	CAGGATTC 359
Db	535	CCTAGTGAAGAAATACCTAGT	TTGACAGCAAGATG	GATATATCTGTGSCCTTAAC	CGAATTC 594
Qy	360	CCTCGAGGCCTAGTGCAA	CCCTAAGAA	GACAGAGTTCAGGGCCTGTGGTGTACTTCG	CAGCAT 419
Db	595	CCTCATGGGAGGAAGAACCA	TTCTGACAAAGGT	CCCCAAACGAGTTGTTCTTC	TGCAACAT 654
Qy	420	GGCCTAGTTGGAGTGTCTAG	CAACTGGATTTCCAA	CCCTGCGCCAAACATAGCCTCGGGCTTC	479
Db	655	GGCTTGCTGGCAGATTC	TAGTAACTGGGT	CACAAACCTTCGCCACACGAC	CCCTGGGCTTC 714
Qy	480	ATTCTGGCAGATGCTGGT	TTTGACGCTGTGGATGGG	GAACAGCAGGGGAAACGCTCGT	CT 539
Db	715	ATTCCTGCTGATGCTGGT	TTTGACGCTGTGGATGGG	CAACAGCAGAGGAAATAC	CTCGTCT 774
Qy	540	CGAAAA	CACAAGACACTCTCCATAG	ACCAAGATGAGTTCCTGGGCTTC	CAGTATGATGAG 599
Db	775	CGGAAA	CATAAGACACTCTAGTTCT	CAGATGAAATCTGGGCTTC	CAGTATGATGAG 834
Qy	600	ATGCTAGTGTGTGACCTTC	CTGCAAGTAAACCTTTAT	TTTTCAGAAAAACGGGCCAGGAA	659
Db	835	ATGGCAAAATATGACCTAC	CAGCTTCCATTAACCTTCAT	CTCGAATAAAAC	CTGGCCAGAA 894
Qy	660	AAGATCTATATGTCGCTAT	TCACAGGCACCACTGGCTTTA	TGTGCATTTTCCACC	719
Db	895	CAAGTGATTAATGTGGT	CTCTCAAGGCACCACTATAG	TTTATAGCATTTTCACAG	954
Qy	720	ATGCCAGAGCTGGCTCAGA	MAATCAAAATGATTTTCTTTAG	CACCATAGCCACTGTT	779
Db	955	ATCCCTGAGCTGGCTAAA	AGGATTAATAATGTTTTTGC	CCCTGGGCTCTGTGGCTTC	CGTC 1014
Qy	780	AAGATGCAAAAGCCCGG	ACCAATTTTGTGCTGCAGATATGATGATCA	AGGGA	839
Db	1015	GCCTTCGTACTAGCCCTAT	GGCCAAATAGGACGATTA	CGAGATCACTCATTAAGGAC	1074
Qy	840	TTGTTTGGCAAAAAGAAAT	TTCTGTATCAGACACAGATTTCTCAGACAACT	TGTTATTTAC	899
Db	1075	TTATTTGGAGACAAGAAT	TTCTTTCCACAGATCGGTTTTTGA	GTGGCTGGGTACCCAC	1134
Qy	900	CTTTGTGGCCAGGTGATCT	TGTGATCAGATTTGTATGATATATCAT	TGTTACTTTCTGGGTGGA	959
Db	1135	GTTTTGCACTCATGT	CATCTGAAGGAGCTCTGTGGAAATCTCTGTTTCTCTGTGTGGA	1194	
Qy	960	TTCAACACCAACATATGA	ACATAGCCGAGCAAGTGTATATGCTGCCACACCTCTTGCT	1019	
Db	1195	TTTAAATGAGAGAAAT	TTAAATATGTCTAGGTGATGTA	TATACAAACATTTCTCTGCT	1254
Qy	1020	GGAA	CATCTGTGCAAAATTTCTTACATGGAGCGAGGAGTGAATTTCTGTGA	ACTCCGG 1079	
Db	1255	GGAACTTCTGTGCAAAA	CATGTTTACCTGGAGCCAGGCTGTAAATTTCCAAA	AGTTTCAA 1314	
Qy	1080	GCATTTGACTGGGGAGTG	AGACCAAAATCTGGAAAATTCGATATCAGCCACCTCCGTGA	1139	
Db	1315	GCCTTTTGACTGGGAG	CAGTGGCCAAAGATTTATTTTCAATTAACAC	CAGAGTTATCTCC 1374	
Qy	1140	AGGTACAGAGTCAGAGAT	TATGACCGTCCCTACAGCAATGTGGACAGGAGGT	CTCAGACTGG 1199	
Db	1375	ACATACATGTGAAGGAC	ATGCTTGTGCCACTCGACTCTGGAGCGGGGT	CTCAGACTGG 1434	
Qy	1200	CTTTCAATCCAGAAGAG	GTGAATGCTGCTCTGAGGTGACCAACCTCATCTAC	CAT 1259	
Db	1435	CTTG	CAGATGTCTACGAGCTCAATCTTACTGACTCAGATTCACCA	ACTTGGTGTTCAT 1494	
Qy	1260	AAGAATATTTCTG	AATGGGCTCAGCTGGATTTTATCTGGGGTTTGGATGCTCTCACCGT	1319	
Db	1495	GAGAGCATTC	CGAATGGGACATCTTGACTTCAITTTGGGSCCTGGATGCCCTTGAGG	1554	
Qy	1320	ATGTACAAATGAAT	TCATCCATCTGATGCAGCAGGAGGACCA	ACCT 1366	
Db	1555	CTTTATATAAAATTTAA	TCTAAATGAGGAAATATCAGTGA	AGACT 1601	

RESULT 13

US-10-108-260A-1262
; Sequence 1262, Application US/10108260A

```

; Publication No. US20040005560A1
;
; GENERAL INFORMATION:
;
; APPLICANT: HELIX RESEARCH INSTITUTE
;
; TITLE OF INVENTION: NO. US20040005560A1el full length cDNA
;
; FILE REFERENCE: HI-A0106
;
; CURRENT APPLICATION NUMBER: US/10/108,260A
;
; CURRENT FILING DATE: 2002-03-27
;
; NUMBER OF SEQ ID NOS: 5458
;
; SOFTWARE: PatentIn Ver. 2.1
;
; SEQ ID NO 1262
;
; LENGTH: 2506
;
; TYPE: DNA
;
; ORGANISM: Homo sapiens
;
; US-10-108-260A-1262

```

Query Match	22.6%;	Score 462.2;	DB 12;	Length 2506;
Best Local Similarity	67.0%;	Pred. No. 1.1e-115;		
Matches 656;	Conservative 0;	Mismatches 323;	Indels 0;	Gaps 0;
QY	388	CAGGTTCCAGGCGCTGTGCTGTTACTGTCAGCATGSCCTAGTTGGAGGTGCTAGCAACTGGGA	447	
Db	280	CTGCTCCAAACACAGTTGTCTTCTGTCGAAATGGCTCGCTGGCAGATCTCTAGTAACCTGGG	339	
QY	448	TTTCCAAACCTGCCCAACAATAGCTGGGCTTCATTCTGGCAGATGCTGGTTTTGACGCTGT	507	
Db	340	TCACAAACCTTGCCCAACAGCAGCGCTGGGCTTCATTCTTGCTGATGCTGGTTTTGACGCTGT	399	
QY	508	GGATGGGGAACAGCAGGGGGAAAAGCCCTGGTCTCGAATAACACAGACACTCTCCATAGACC	567	
Db	400	GGATGGGCAACAGCAGAGGAAATACCTGGTCTCGGAAACATAAGACACTCTCAGTTTCTC	459	
QY	568	AAGATGAGTTCTGGGCTTTCAGTTATCATGATGGCTAGGTTTGACTTTCCTGTCAGTGA	627	
Db	460	AGGATGAATTCTGGGCTTTCAGTTATGATGAGATGGCAAAATATGACTACCAAGTTCCA	519	
QY	628	TAAACTTTATTTTGCAGAAAACGGGCGAGAAAAGATCTATTATGTGCGGTATTTCACAGG	687	
Db	520	TTAACTTCAATCTGAATAAAACTGGCCAGAAACAAGTGATATTATGTGGTCAITCTCAAG	579	
QY	688	GCACCACATGGGCTTTTATTGCATTTTCCACCATGCCAGAGCTGGCTCAGAAATCAAAA	747	
Db	580	GCACCACATAGGTTTTATAGCATTTTTCACAGATCCCTGAGCTGGCTTAAAGGATTAATA	638	
QY	748	TGATTTTGTCTTTAGACCCATAGCCACTGTTAAGCATGCAAAAAGCCCGGACCAAAAT	807	
Db	640	TGTTTTTGGCCCTGGGTCCTGTGGCTTCGGTGGCTTCTGTACTAGCCCTATGGCCAAAT	699	
QY	808	TTTTTGTGTGCCAGATATAGATGATCAAGGGAATGTTTGGCAAAAAGAAATTTCTGTATC	867	
Db	700	TAGACGATTAACAGATCATCTCATTAAGACATTATTGGAGACAAAGAAATTTCTTCCCC	759	
QY	868	AGACGATTTCTCAGACAACTTGTTATTACCTTTGTGGCCAGGTGATTTCTTGATCAGA	927	
Db	760	AGAGTGGCTTTTGAAGTGGCTGGGTACCCACGTTTTGCACTCATGTCAATCTGAAGAGC	819	
QY	928	TTTCTAGTAATATCATGCTTACTCTGGGTGGATTCACACCAACAATATCAACATGAGCC	987	
Db	820	TCCTGGAAATCTCTGTTTCTCTGTGTGGATTTAATGAGAGAAATTTAAATATGCTTA	879	
QY	988	GAGCAATGTATATGTGCGCCCAACTCTTGTGTGAAACATCTGTGCCAAAAATTTCTACACT	1047	
Db	880	GAGTGGATGTATATACAAACAATTTCTCTGTGTGAACTTCTGTGCCAAAAATGTTTACACT	939	
QY	1048	GGAGCCAGGAGTGAAATTTCTGGTGAACTCCGGGCAATTTGACTGGGGAGTGAGACCAAAA	1107	
Db	940	GGAGCCAGGCTGTTAAATTCCAAAGTTTCAAGCCTTTGACTGGGGAGCAGTGGCCAAAGA	999	
QY	1108	ATCTGGAAAAATGCAATCAGCCAACTCCTGTGAAGGTACAGAGTACAGAGATATGACGGTCC	1167	

Db 1000 ATTATTTTCAATACACAGAGTTATCTCCAGTACAAATGTGAGACATGCTTGGC 1059
 Qy 1156 CTACAGCAATGTGACAGAGGTGAGTCTGGCTTTTCAATTCAGAGAGAGCTGAAATGC 1227
 Db 1060 CGACTGCACTGTGAGCGGGGTGAGTCTGGCTTGCAGATGTCTACGAGCTCAATATCT 1119
 Qy 1228 TGCTCTGAGGTGACCAACCTCATCTACCAATAGAAATATTCCTGAATGGCTCAGCTGG 1287
 Db 1120 TACTGACTCAGATCACCAACTTGGTGTCCATGAGAGCAATTCGGAAATGGAGCATTTG 1179
 Qy 1288 ATTTCATCTGGGTTTGGATGCTCTCACCCTGATGTACAAATGAATCATCTGATGC 1347
 Db 1180 ACTTCATTTGGGCTGGATGCCCTTGGAGGCTTTATAATAAAATTAATTAATCTAATGA 1239
 Qy 1348 AGCAGGAGGAGACCAACCT 1366
 Db 1240 GGAATATCATGTGAAGCT 1258

RESULT 14

US-10-420-564-3
 ; Sequence 3, Application US/10420564
 ; Publication No. US20040001819A1
 ; GENERAL INFORMATION:
 ; APPLICANT: Bolen, Paul L.
 ; APPLICANT: Cihak, Paul, L.
 ; TITLE OF INVENTION: Recombinant Kid Pregastric Esterase and Methods for
 ; TITLE OF INVENTION: its
 ; TITLE OF INVENTION: Production and Use
 ; FILE REFERENCE: IPR-0009
 ; CURRENT APPLICATION NUMBER: US/10/420,564
 ; PRIOR FILING DATE: 2003-04-22
 ; PRIOR APPLICATION NUMBER: US/10/043,665B
 ; PRIOR FILING DATE: 2002-09-24
 ; PRIOR APPLICATION NUMBER: US 09/186,489
 ; PRIOR FILING DATE: 1998-11-05
 ; NUMBER OF SEQ ID NOS: 7
 ; SOFTWARE: Patentin ver. 2.0
 ; SEQ ID NO 3
 ; LENGTH: 1411
 ; TYPE: DNA
 ; ORGANISM: Kid (Goat)
 US-10-420-564-3

Query Match 22.3%; Score 456.8; DB 12; Length 1411;

Best Local Similarity 62.0%; Pred. No. 2.2e-114;

Matches 741; Conservative 0; Mismatches 452; Indels 3; Gaps 1;

Qy 167 AATGAAATGTGGCTTCTGATTCTGGTGGCTATATGTTCCAGAGAAATGTGAATTCAGT 226
 Db 61 ATTGAATGTGGCTACTTGTAAACGGTGTGTTTCAATCCACATGTCTGAAATGCATT 120
 Qy 227 ACATATGC---CACTAAAGCTGTGACCCAGAGCATTCATGAATATTAAGTGAATCAT 283
 Db 121 TTGTTTCTTGGAAAAATTTGTAAGAACCTTGAAGCCAGTATGAATGTGAGTCAGATGAT 180
 Qy 284 CCAACATCAAGCTATCTCTGTGAGATATGAATGCTGCACTGCAATGAGTGGTATATCCT 343
 Db 181 TTCCTTCTGGGCTACCAAGTGAATGCAATGAATGATATTAATGAGTGGTATATATCCT 240
 Qy 344 TTCTTTTAAAGGATTTCTTCGAGGCTTAGTGCACCTAAAGAGACAGGTTCCAGGCTGT 403
 Db 241 TCAGGTCTATCGAATTCCTCTGAAAGAAATGATGCTAATCATTTAGTCTCAGAGCTGT 300
 Qy 404 GTGTGACTGAGCATGCGCTAGTTGGAGGTGCTAGCACTGGAATTTCAACCTGCCAA 463
 Db 301 TGTGTTTCTGAGCATGTGCTTCTTGCCTCAGCTACAACTGGATTTCCAACTTCCAA 360
 Qy 464 CAATAGCCTGGGCTTCATTTCTGGCAGATGCTGGTGTGACGTGTGGATGGGAGACAG 523
 Db 361 CAACAGCCTGGGCTTCCTCTGGCAGATGCTGTTATGACGTGTGGCTGGGAGACAG 420

Qy 524 GGGAAAGCGCTGGTCTCGAAAAACAAGACACTCTCCATAGACCAAGATGAGTTCTGGGC 583
 Db 421 AGAAAAACCTTTGGGCCCCAGGAAACATTTATATTCACAGACTCCCTCGAATTTCTGGC 480
 Qy 584 TTTCACTATGATGAGATGGCTTGGCTTGGCTTGGCTTGGCTTGGCTTGGCTTGGCTTGGC 643
 Db 481 TTTCACTTGTATGAAATGGCTGAATATGACCTTCCATCTACAATGATTTCACTTAA 540
 Qy 644 GAAACGGCCAGGAAAGATCTATATGTGGCTTATTCACAGGACACCATGGCTT 703
 Db 541 GAGAACAGACAGAGAGACTACATATGTTGGCCATTTCCCAAGGACACCATGGCTT 600
 Qy 704 TATTGCAATTTCCACCATGCGAGAGTGGCTCAGAAAAATCAAAATGATTTTGTCTTAGC 763
 Db 601 TGTGGCTTTTCTACCAATCCCACTGGCTGAAATAATCGAGTCTTCCATGCAATAGC 660
 Qy 764 ACCATAGCCACTGTTAAGCATGCAAAAAGCCCCGGGACCAAAATTTTGTCTGCCAGA 823
 Db 661 CCCAGTGGCCACAGTGAAGCACAACCCAGAGCTGTTTAAACAACTTGCATTTATTCCTCA 720
 Qy 824 TATGATGATCAAGGATGTTTGGCAAAAAGAAATTTCTGATACAGACCAAGATTTCTCAG 883
 Db 721 CTCTCTTCAAGATTAATTTGGTAAACAAATGTTCTACCCACAAATTTTGTGAACA 780
 Qy 884 ACACTGTTTATTTACCTTTGTGGCCAGGTGATTCTTGTATCAGATTTGTAGTAATCAT 943
 Db 781 ATTCTTGTGTTGAAGTGTGCTCTCTGAGACACTGGATGCTCTTGTGAAGATGCTT 840
 Qy 944 GTTACTTCTGGGTGGATTCACCAACCAATATGAATGACGCGCAGCAAGTGTATATGC 1003
 Db 841 GTTGGCAATCTAGTGGAGCTGCAATAAATACTTCAACATGAGTCTGCTTAGATGTATGT 900
 Qy 1004 TGCCCACTCTTGTGGAAACATCTGTGCAAAATATTTCTACACTGGAGCCAGGCAAGTAA 1063
 Db 901 AGCACAATATCCAGCAGGAGCTTCTGTTCAAAACATCTCCACTGGAGACAGGCTATTAA 960
 Qy 1064 TTCTGTGAACTCCGGGCAATTTGACTGGGGAGTGAGACCAAAATCTGGAATAATGAA 1123
 Db 961 GTCTGGGAAATTTCCAAGCTTTTGACTGGGGAGCCCTCAGTTGAGAACCTTAATGCAATTA 1020
 Qy 1124 TCAGCCAACTCTGTAAAGTACAGAGTACAGATATGACGGTCCCTACAGCAATGTGAC 1183
 Db 1021 TCAGCCCACTCTCCATCTACAATTTAACGCCATGATGTCCCAATTCAGATGGAG 1080
 Qy 1184 AGAGTCTAGGACTGGCTTTCAATTCAGAGAGAGCTGAAATGCTGCTCTGAGGTGAC 1243
 Db 1081 TGTGGCCAGACCTGTTGGCTGACCCCTCAGGATGTTGACCTTTTGTCTTCAAACTCTC 1140
 Qy 1244 CAACCTCATCTACCAATAAGAAATTTCTGAATGGGCTCAGTGGGATTTTCACTCTGGGTT 1303
 Db 1141 TAACTCATCTACCAAGGAAATTTCCAAATTAATCATCTGAGCTTTATCTGGGCAAT 1200
 Qy 1304 GGATGCTCTCACCCTATGATGATGAATCAATCATCTGATGACAGGAGGAGA 1359
 Db 1201 GGATGCACTCAAGAGTTTACAATGAAATTTATTTCTTGTGGCAAAAGACAAA 1256

RESULT 15

US-10-312-088-7

; Sequence 7, Application US/10312088

; Publication No. US20030219862A1

; GENERAL INFORMATION:

; APPLICANT: Agarwal, Pankaj

; APPLICANT: Cogswell, John P.

; APPLICANT: Kabnic, Karen S.

; APPLICANT: Lai, Ying-Ta

; APPLICANT: Martensen, Shelby A.

; APPLICANT: Murdock, Paul R.

; APPLICANT: Smith, Randall F.

; APPLICANT: Strum, Jay C.

; APPLICANT: Xiang, Zhaoying

; APPLICANT: Xie, Qing

```

; APPLICANT: Rizni, Safia K.
; TITLE OF INVENTION: NOVEL COMPOUNDS
; FILE REFERENCE: GP50029
; CURRENT APPLICATION NUMBER: US/10/312,088
; CURRENT FILING DATE: 2002-12-20
; PRIOR APPLICATION NUMBER: PCT/US01/19929
; PRIOR FILING DATE: 2001-06-22
; PRIOR APPLICATION NUMBER: 60/213,161
; PRIOR FILING DATE: 2000-06-22
; PRIOR APPLICATION NUMBER: 60/213,156
; PRIOR FILING DATE: 2000-06-22
; NUMBER OF SEQ ID NOS: 44
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 7
; LENGTH: 1194
; TYPE: DNA
; ORGANISM: Homo sapiens
US-10-312-088-7

Query Match      22.2%; Score 454.2; DB 13; Length 1194;
Best Local Similarity 62.4%; Pred. No. 1e-113;
Matches 711; Conservative 0; Mismatches 428; Indels 0; Gaps 0;

QY 211 GAAATGTGAATTCAGTACATATGCCAACTAAAGCTGTGGACCCAGAGCAATTCATGAATA 270
Db      |||
QY 271 TTAGTGAATTCATCCAAATCAAGGCTATCCCTGTGAGGAATATGAAATCGCAATCGAAG 330
Db      |||
QY 104 TTAGCCAGATTAATTTCTTACTGCGGGTTATCCCTTATGAAGATGATGATGTACACAAAAG 163
Db      |||
QY 331 ATGGGTATATCCCTTCTGTGTAAACAGGATTCCTCGAGGCTAGTGCACACCTTAAGAAAGACAG 390
Db      |||
QY 164 ATGGTTATATCCCTTGGAAATTTATAGATTTCCATGGAAGAGATGCGCCAGGAGACAG 223
Db      |||
QY 391 GTTCCAGGCTGTGGTGTGTTACTGACAGATGGCTAGTGTGAGGTGCTAGCAACTGGATTT 450
Db      |||
QY 224 CTCGAAAGCTGTGTGTATTTGACAGCATGGCTTAATTTGATCTGCGAGTAACTGGATTT 283
Db      |||
QY 451 CCACCTGCCCCAATAGCTGGCTTCAATCTGGCAGATGCTGGTTTGGACGTGTGA 510
Db      |||
QY 284 GCACCTTGGCCCAACACAGTTGGCTTTCCCTTGGCAGATGATGTTTATGACGTGTGT 343
Db      |||
QY 511 TGGGGAACAGCAGGGGAAACGCTGTGTCTGAAACACAGACACTCTCCATAGACCAAG 570
Db      |||
QY 344 TGGGAACAGCCGAGGAAACACTTGTGTCCAGAAACACCTTAAATTTGTCAACGAAATCAC 403
Db      |||
QY 571 ATGAGTCTGGCTTTTCAGTTATGATGATGATGGCTAGTGTGACCTTCTGCGAGTGATA 630
Db      |||
QY 404 CGGAATACTGGGCTTTTCAGTTTGGATGATGATGCTAAATGATGACCTTCCAGCCACATCA 463
Db      |||
QY 631 ACTTTATTTTGCAGAAACCGGGCCAGGAAAGATCTTATTATGTCGGCTATTTCACAGGGCA 690
Db      |||
QY 464 ATTTTATCATAGAGAAACCTGGACAGAGCGACTTACTACGTGGGCCACTCACAAGGCA 523
Db      |||
QY 691 CCACCATGGGCTTTATGCAATTTTCCACATGCGAGAGTGGCTCAGAAATCAAATGT 750
Db      |||
QY 524 CCACCATAGCTTTTATAGCAATTTTCTACAAACCCAGACCTGGCTTAAAGATTAAGATAT 583
Db      |||
QY 751 ATTTTGTCTTTAGCACCCATAGCACTGTTAAGCATGCAAAAGCCCGGGGACCAATTTT 810
Db      |||
QY 584 TTTTTCACCTGGCTCCAGTTGTTCACAGTTTAAATACACCCCAAGTCTTATGAAAAAATAA 643
Db      |||
QY 811 TGTGTGCGCAGATATGATGATCAGGGATGTTTGGCAAAAGAAATTTCTGTATCAGA 870
Db      |||
QY 644 CAACCCCTTCCAGGCGAGTAGTTAAGGTGTGTTTGGTGAACAAATGTTCCACCCCTCATA 703
Db      |||
QY 871 CCAGATTTCTCAGACCACTTGTATTATTTACCTTTTGGCCAGAGTGAATTTCTGTATCAGATTT 930
Db      |||
QY 704 CATTTGTTTGACCAATTCATTGCCACCAAGTGTGCAATCGAAGCTATTTCCGTCGTATTT 763
Db      |||
QY 931 GTAGTAATATCATGTTACTTCTGGGTGGATTTCAACACCAATATATGAACATGAGCCGAG 990
Db      |||

```

```

Db 764 GCAGCAACTTCTCTATTCTCTGAGTGGATTTGATCCGCAAACTTTAAATATGAGTCGCT 823
QY 991 CRAAGTGTATATGCTGCCACACTCTTGTCTGGGAACATCTCTGCAAAATATTTCTACACTGA 1050
Db 824 TGGATGTTTATTGTCACACATCTCTGCGGGAACATCTGTTTCAAAATATGCTGCACTGGG 883
QY 1051 GCCAGCAGTGAATTTCTGTGAACTCCGGGCAATTTGACTGGGGGAGTGAGACCAAAATC 1110
Db 884 CTCAGCTTTTACCACCTCTGATGAATTCAGAGCTTATGACTGGGGAATGACGCTGATAATA 943
QY 1111 TGGAAATATGCAATCAGCCAACTCTCTGTAAAGTACAGAGTTCAGAGATATGACGGTCCCTA 1170
Db 944 TGAACATTAACAATCAGAGTCTATCCCTCTATATATGACCTGCTGCCATGAAGTCCCTA 1003
QY 1171 CAGCAATGTGGACAGGAGGTTCAGGACTGGCTTTCAAATCCAGAAGACGTGAAAATGCTGC 1230
Db 1004 CTGCTATTTGGGCTGTGGTGGACATGATGTCCTCTGTAACACCCAGGATGTGGCCAGGATAC 1063
QY 1231 TCTCTGAGGTGACCAACCTCATCTACATGAAGATATTTCTGAATGGGCTCAGGTGGATT 1290
Db 1064 TCCCTCAAATCAAGAGTCTTCTATTCTTTAAGCTATTTGCCAGATTGGAACCACTTTGATT 1123
QY 1291 TCATCTGGGGTTTTGGATGCTCTCACCCTATGTACAAATGAAATCATCCATCTGAATCAG 1349
Db 1124 TTGCTGGGGCCCTCGATGCCCTCAACGGATGTACAGTGAATCATAGCTTTAATGAAG 1182

```

Search completed: February 19, 2004, 14:12:27
Job time : 747.122 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2004 Compugen Ltd.

OM nucleic - nucleic search, using sw model

Run on: February 19, 2004, 04:41:57 ; Search time 4665.38 Seconds
(without alignments)
11127.553 Million cell updates/sec

Title: US-10-042-431-46
Perfect score: 1269
Sequence: 1 atgttgaaacctgtcaag.....gacggtgtgagccgtattg 1269

Scoring table: IDENTITY_NUC
Gapop 10.0 , Gapext 1.0

Searched: 2888711 seqs, 20454813386 residues

Total number of hits satisfying chosen parameters: 5777422

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database :

GenEmbl.*

- 1: gb.ba.*
- 2: gb.htg.*
- 3: gb.in.*
- 4: gb.om.*
- 5: gb.ov.*
- 6: gb.pat.*
- 7: gb.ph.*
- 8: gb.pl.*
- 9: gb.pr.*
- 10: gb.ro.*
- 11: gb.sts.*
- 12: gb.sy.*
- 13: gb.un.*
- 14: gb.vi.*
- 15: em.ba.*
- 16: em.fun.*
- 17: em.hum.*
- 18: em.in.*
- 19: em.mu.*
- 20: em.om.*
- 21: em.or.*
- 22: em.ov.*
- 23: em.pat.*
- 24: em.ph.*
- 25: em.pl.*
- 26: em.ro.*
- 27: em.ste.*
- 28: em.un.*
- 29: em.vi.*
- 30: em.htg_hum.*
- 31: em.htg_inv.*
- 32: em.htg_other.*
- 33: em.htg_mus.*
- 34: em.htg_pln.*
- 35: em.htg_ror.*
- 36: em.htg_man.*
- 37: em.htg_vrt.*
- 38: em.sy.*
- 39: em.htgo_hum.*
- 40: em.htgo_mus.*
- 41: em.htgo_other.*

Pred. No. is the number of results predicted by chance to have a

score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB	ID	Description
1	1132	88.4	1210	6	AX521798	AX521798 Sequence
2	1104.8	87.1	1272	6	AX477670	AX477670 Sequence
3	957	75.4	2577	10	BC031933	BC031933 Mus muscu
4	752	59.3	1116	6	AX521801	AX521801 Sequence
5	528.2	41.6	2582	9	AK091558	AK091558 Homo sapi
6	527.8	41.6	2481	6	AX411234	AX411234 Sequence
7	527.8	41.6	2481	9	HSU08464	HSU08464 Human lysos
8	527.8	41.6	2481	9	HSU08464	HSU08464 Human lysos
9	527.8	41.6	2481	9	HSU08464	HSU08464 Human lysos
10	526.2	41.5	2493	6	BD094072	BD094072 Shear str
11	526.2	41.5	2493	9	HUMLIPCHL	HUMLIPCHL
12	524.6	41.3	2586	9	BC012287	BC012287 Homo sapi
13	489.2	38.6	1378	6	A26689	A26689 Precursor o
14	489.2	38.6	1378	6	A26690	A26690 Precursor o
15	470	37.0	1365	9	HSGLR	HSGLR
16	470	37.0	1367	6	A01046	A01046 H.sapiens m
17	470	37.0	1367	6	A012714	A012714 Pregartric
18	467.6	36.8	1336	6	A01157	A01157 R.norvegicu
19	467.6	36.8	1336	6	A01157	A01157 R.norvegicu
20	462.2	36.4	2506	9	AK096406	AK096406 Homo sapi
21	461	36.3	2358	10	MMYACLY	MMYACLY
22	459.8	36.2	3144	10	S81497	S81497 lysosomal a
23	458	36.1	1417	4	BOVME	BOVME
24	454	35.8	1228	6	AX392821	AX392821 Sequence
25	448.4	35.3	1473	6	AX392824	AX392824 Sequence
26	446.8	35.2	1195	6	AX472752	AX472752 Sequence
27	444.2	35.0	1137	6	AR039022	AR039022 Sequence
28	444.2	35.0	1137	6	AR0392633	AR0392633 Sequence
29	444.2	35.0	1140	6	AR039023	AR039023 Sequence
30	444.2	35.0	1140	6	AR0392634	AR0392634 Sequence
31	444.2	35.0	1146	6	AR039024	AR039024 Sequence
32	444.2	35.0	1146	6	AR0392635	AR0392635 Sequence
33	444.2	35.0	1198	6	A57760	A57760 Sequence 5
34	444.2	35.0	1528	6	A57756	A57756 Sequence 1
35	444.2	35.0	1528	6	AR039021	AR039021 Sequence
36	444.2	35.0	1528	6	AR0392632	AR0392632 Sequence
37	444.2	35.0	1651	4	CFLIPASE	CFLIPASE
38	441.8	34.8	1800	5	BC052131	BC052131 Danio rer
39	441	34.8	1137	6	A39301	A39301 Sequence 2
40	441	34.8	1137	6	A39303	A39303 Sequence 4
41	441	34.8	1146	6	A39305	A39305 Sequence 6
42	441	34.8	1146	6	A39300	A39300 Sequence 1
43	424.8	33.5	1260	6	AX645003	AX645003 Sequence
44	419.8	33.1	1086	6	AX710607	AX710607 Sequence
45	388.8	30.6	1048	6	A57758	A57758 Sequence 3

ALIGNMENTS

RESULT 1	AX521798	Sequence 1 from Patent WO0236731.	1210 bp	DNA	linear	PAT 05-OCT-2002
LOCUS	AX521798	Sequence 1 from Patent WO0236731.	1210 bp	DNA	linear	PAT 05-OCT-2002
DEFINITION	AX521798	Sequence 1 from Patent WO0236731.	1210 bp	DNA	linear	PAT 05-OCT-2002
ACCESSION	AX521798	Sequence 1 from Patent WO0236731.	1210 bp	DNA	linear	PAT 05-OCT-2002
VERSION	AX521798.1	GI:23572854	1210 bp	DNA	linear	PAT 05-OCT-2002
KEYWORDS	Homo sapiens (human)					
SOURCE	Homo sapiens (human)					
ORGANISM	Homo sapiens					
REFERENCE	Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;					
AUTHORS	Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.					
TITLE	1					
JOURNAL	Xiao, Y.					
	Human lysosomal acid lipase					
	Patent: WO 0236731-A 1 10-MAY-2002;					

FEATURES		Bayer Aktiengesellschaft (DE)	
source		Location/Qualifiers	
		1..1210	
		/organism="Homo sapiens"	
		/mol_type="genomic DNA"	
		/db_xref="taxon:9606"	
CDS		1..1197	
		/note="unnamed protein product"	
		/codon_start=1	
		/protein_id="CAD52970.1"	
		/db_xref="GI:23572855"	
		/translation="MELKLLFRDVLFWAEMTRPLDKOSEIHHQSYFCEEEVEVATED GYLISVNRIPRGVQPKTGSRPVLLQHLVGGASNWIISNPNLSLFIADAGFDV WMNSGNASWRKHTLSIDQBFWAFSYDENARFDPVINFILKQTKQEKIYYVG SQGTWGFAPSMPELAOKIKMYPALAPIATVHAKSPGTRFLLPDMWIKGLPGKK EFYCTPREFRLQVLYLCOGVILDOICSNIMLLLGGPNTNNMNSRASVAAHTLAGTS VQNLHWSQAVNSGEHLRAEDKSGEYLNLEKCNQPPFVRVREDMTVPVAAHTGGQDL SNPDVNLLESEVTNLIYHNKIPFWAHVDFIWLGLDAPHRWNEIHLKQOEBETNUSQG RCEAVL"	
BASE COUNT		338 a	296 g 312 t
ORIGIN			
		Query Match 88.4%; Score 1122; DB 6; Length 1210;	
		Best Local Similarity 100.0%; Pred. No. 0;	
		Matches 112; Conservative 0; Mismatches 0; Indels 0; Gaps 0;	
Qy	148	AGTGAATATCCCAACATCAAGGCTATCCCTGTGAGGAATATGAGTGCACAACTGAAGAT	207
Db	73	AGTGAATATCCCAACATCAAGGCTATCCCTGTGAGGAATATGAGTGCACAACTGAAGAT	132
Qy	208	GGGTATATCTTTCTGTATACAGGATTCCTCGAGGCTAGTGCACCTAAGAGACAGCT	267
Db	133	GGGTATATCTTTCTGTATACAGGATTCCTCGAGGCTAGTGCACCTAAGAGACAGCT	192
Qy	268	TCCAGGCTGTGGTGTACTGCAGCATGGCTAGTGTGGAGGTGTGACCACTGATTTCC	327
Db	193	TCCAGGCTGTGGTGTACTGCAGCATGGCTAGTGTGGAGGTGTGACCACTGATTTCC	252
Qy	328	AACCTGCCCAACATAGCTGGGTTCAATCTGGCAGATGCTGTTTACGCTGTGGATG	387
Db	253	AACCTGCCCAACATAGCTGGGTTCAATCTGGCAGATGCTGTTTACGCTGTGGATG	312
Qy	388	GGGAACAGCAGGGGAAACGCTGTCTCGAAACACAAAGACATCTCCATAGACCAAGAT	447
Db	313	GGGAACAGCAGGGGAAACGCTGTCTCGAAACACAAAGACATCTCCATAGACCAAGAT	372
Qy	448	GAGTTCTGGGCTTTTCACTATATGATGAGATGGCTAGGTTTCACTTCTGTCAGTGTAAAC	507
Db	373	GAGTTCTGGGCTTTTCACTATATGATGAGATGGCTAGGTTTCACTTCTGTCAGTGTAAAC	432
Qy	508	TTTATTTTGCAGAAACAGGGCCAGGAAAGATCTATTATGCTCGCTATTCCACAGGCGACC	567
Db	433	TTTATTTTGCAGAAACAGGGCCAGGAAAGATCTATTATGCTCGCTATTCCAGGCGACC	492
Qy	568	ACCATGGGCTTTATTTGATTTTCCATGCTCCAGAGCTGGCTCAGAAAATCAAAATGTAT	627
Db	493	ACCATGGGCTTTATTTGATTTTCCATGCTCCAGAGCTGGCTCAGAAAATCAAAATGTAT	552
Qy	628	TTTCTTTTACACCATAGCACCTGTTTAAAGATGCAAAAGCCCGGGACCAATTTTG	687
Db	553	TTTCTTTTACACCATAGCACCTGTTTAAAGATGCAAAAGCCCGGGACCAATTTTG	612
Qy	688	TTGCTGCCAGATATGATGATCAAGGGATTGTTTGGCAAAAAGAAATTTCTGTATCAGACC	747
Db	613	TTGCTGCCAGATATGATGATCAAGGGATTGTTTGGCAAAAAGAAATTTCTGTATCAGACC	672
Qy	748	AGATTCTCAGACAACTGTTTATTTACTTTTGGCCAGGATGATCTTTGATCAGATTGT	807
Db	673	AGATTCTCAGACAACTGTTTATTTACTTTTGGCCAGGATGATCTTTGATCAGATTGT	732
Qy	808	AGTAATATCATGTTACTTCTGGGTGGATTCAACACCAACATATGAACATGAGCCGACGA	867
Db	733	AGTAATATCATGTTACTTCTGGGTGGATTCAACACCAACATATGAACATGAGCCGACGA	792

Qy	868	AGTGTATATGTCGCCACACACTCTTGTGGAAACATCTGTGCARAAATATTTTACACTGGAGC	927
Db	793	AGTGTATATGTCGCCACACACTCTTGTGGAAACATCTGTGCARAAATATTTTACACTGGAGC	852
Qy	928	CAGCAGTGAATTTCTGGTGAATCCCGGGCAATTTGACATGGGGAGTGAGACCAAAAATCTG	987
Db	853	CAGCAGTGAATTTCTGGTGAATCCCGGGCAATTTGACATGGGGAGTGAGACCAAAAATCTG	912
Qy	988	GAAAAATGCAATCAGCCAACTCCTGTAAGGTACAGAGTACAGATATGACGGTCCCTACA	1047
Db	913	GAAAAATGCAATCAGCCAACTCCTGTAAGGTACAGAGTACAGATATGACGGTCCCTACA	972
Qy	1048	GCAATGTGGACAGAGAGTTCAGGACTGGCTTTTCAAAATCCAGACAGCTGAAAAATCTGCTC	1107
Db	973	GCAATGTGGACAGAGAGTTCAGGACTGGCTTTTCAAAATCCAGACAGCTGAAAAATCTGCTC	1032
Qy	1108	TCTGAGGTGACCAACCTCATCTACCAATAAGAAATATTCCTGAATGGGCTCAGTGGATTTC	1167
Db	1033	TCTGAGGTGACCAACCTCATCTACCAATAAGAAATATTCCTGAATGGGCTCAGTGGATTTC	1092
Qy	1168	ATCTGGGGTTTGGATGCTCTCTCACCGTATGTACAAATGAAATCATCTGATGCAGCAG	1227
Db	1093	ATCTGGGGTTTGGATGCTCTCTCACCGTATGTACAAATGAAATCATCTGATGCAGCAG	1152
Qy	1228	GAGGAGACCAACCTTTTCCAGGAGCGGTGTGAGGCCGTATTG	1269
Db	1153	GAGGAGACCAACCTTTTCCAGGAGCGGTGTGAGGCCGTATTG	1194
RESULT 2			
AX477670			
LOCUS			
DEFINITION			
ACCESSION			
VERSION			
KEYWORDS			
SOURCE			
ORGANISM			
REFERENCE			
AUTHORS			
TITLE			
JOURNAL			
FEATURES			
source			
BASE COUNT			
ORIGIN			
Query Match 87.1%; Score 1104.8; DB 6; Length 1272;			
Best Local Similarity 97.2%; Pred. No. 0;			
Matches 1143; Conservative 0; Mismatches 12; Indels 21; Gaps 1;			
Qy	115	AAAGCTGTGGAGCCCAAGAGCATTCATGAAATATTAGTGAATATCCCTTTCTGTAAACAGGATT	174
Db	94	AATGAAGTGAATTCCTCTGAGGTGTGGATGAATATCTAGTGAATATCCCAACATCAAGGCTAT	153
Qy	175	CCCTGTGAGGAAATGAAAGTGCACACTGAAAGTGGGTATATCCCTTTCTGTAAACAGGATT	234
Db	154	CCCTGTGAGGAAATGAAAGTGCACACTGAAAGTGGGTATATCCCTTTCTGTAAACAGGATT	213
Qy	235	CCTCAGGCTGTAGTGCACCTTAAGAGACAGGTTCCAGGCTGTGGTGTACTGCAGCAT	294
Db	214	CCTCAGGCTGTAGTGCACCTTAAGAGACAGGTTCCAGGCTGTGGTGTACTGCAGCAT	273

QY 295 GGCCTAGTTGAGGTGTAGCACTGATTTCCAACTGCCACATAGCCTGGGCTTC 354
 DB 274 GGCCTAGTTGAGGTGTAGCACTGATTTCCAACTGCCACATAGCCTGGGCTTC 333
 QY 355 ATTCTGGCAGATGCTGTTTGGCTGATGGGGAACGAGGGAACGCTGCTCT 414
 DB 334 ATTCTGGCAGATGCTGTTTGGCTGATGGGGAACGAGGGAACGCTGCTCT 393
 QY 415 CGAAACACACAGACACTCTCCATAGACCAAGATGATGCTGGGCTTTCAGTTATGATGAG 474
 DB 394 CGAAACACACAGACACTCTCCATAGACCAAGATGATGCTGGGCTTTCAGTTATGATGAG 453
 QY 475 ATGGCTAGTTGAGCTTCTGCTGATGATAAATTTTATTTGCAAGAAACGGGCGAGGAA 534
 DB 454 ATGGCTAGTTGAGCTTCTGCTGATGATAAATTTTATTTGCAAGAAACGGGCGAGGAA 513
 QY 535 AAGATCTATTATGCTGGCTATTCACAGGCGACCACTGCTGGCTTATGATTTCCACC 594
 DB 514 AAGATCTATTATGCTGGCTATTCACAGGCGACCACTGCTGGCTTATGATTTCCACC 573
 QY 595 ATGCCAGAGCTGGCTCAGAAATCAAAATGATTTTGTCTTAGCACCATAGCCACTGTT 654
 DB 574 ATGCCAGAGCTGGCTCAGAAATCAAAATGATTTTGTCTTAGCACCATAGCCACTGTT 633
 QY 655 AAGCATGCAAAAGCCCGGACCAAAATTTTGTCTGCGAGATATGATCAAGGGA 714
 DB 634 AAGCATGCAAAAGCCCGGACCAAAATTTTGTCTGCGAGATATGATCAAGGGA 693
 QY 715 TTGTTTGGCAAAAGAAATTTCTGATCAGACCAAGATTTCTCAGACAACTTGTATTATAC 774
 DB 694 TTGTTTGGCAAAAGAAATTTCTGATCAGACCAAGATTTCTCAGACAACTTGTATTATAC 753
 QY 775 CTTTGTGGCAGGTGATTTCTGATCAGATTTGTAGTAATATCATTTACTTCTGGGTGGA 834
 DB 754 CTTTGTGGCAGGTGATTTCTGATCAGATTTGTAGTAATATCATTTACTTCTGGGTGGA 813
 QY 835 TTCAACACCAACAATATGAACAT-----GAGCCGAGCAAGTGTGA 873
 DB 814 TTCAACACCAACAATATGAACATGATTAATCTGTTGTTACAGAGCGGAGCAAGTGTGA 873
 QY 874 TATGCTGCCACACTCTTCTGCGAACTCTGTGCAAAATATTTCTACACTGAGCCAGGCA 933
 DB 874 TATGCTGCCACACTCTTCTGCGAACTCTGTGCAAAATATTTCTACACTGAGCCAGGCA 933
 QY 934 GTGAATTTCTGTTGAACTCTGGGCAATTTGACTGGGGAGTGAGACCAAAATCTGGAATA 993
 DB 934 GTGAATTTCTGTTGAACTCTGGGCAATTTGACTGGGGAGTGAGACCAAAATCTGGAATA 993
 QY 994 TGCAATCAGCCAACTCTGTTAAGGTACAGAGTCAGAGATATGACCGTCCCTACAGCAATG 1053
 DB 994 TGCAATCAGCCAACTCTGTTAAGGTACAGAGTCAGAGATATGACCGTCCCTACAGCAATG 1053
 QY 1054 TGACAGGAGGTGAGGCTGCTTCAATCCAGAGAGCTGAAATCTGCTCTCTGAG 1113
 DB 1054 TGACAGGAGGTGAGGCTGCTTCAATCCAGAGAGCTGAAATCTGCTCTCTGAG 1113
 QY 1114 GTGACCAACTCATCTACCATAGAAATATTTCTGAATGGCTCAGTGGATTTCTATCTGG 1173
 DB 1114 GTGACCAACTCATCTACCATAGAAATATTTCTGAATGGCTCAGTGGATTTCTATCTGG 1173
 QY 1174 GGTTTGGATGCTCTCAGCGATGATGAAATGAAATCATCTCTGATGACGAGGAGGAG 1233
 DB 1174 GGTTTGGATGCTCTCAGCGATGATGAAATGAAATCATCTCTGATGACGAGGAGGAG 1233
 QY 1234 ACCAACCTTTCCAGGAGCGGTGTGAGCGCGATTG 1269
 DB 1234 ACCAACCTTTCCAGGAGCGGTGTGAGCGCGATTG 1269

RESULT 3
 BC031933 2577 bp mRNA linear ROD 20-SEP-2002
 LOCUS Mus musculus, RIKEN cdna 4632427C23 gene, clone MGC:25790
 DEFINITION

ACCESSION
 VERSION
 KEYWORDS
 SOURCE
 ORGANISM

REFERENCE
 AUTHORS
 TITLE
 JOURNAL

REMARK
 COMMENT

FEATURES
 source

CDS

BASE COUNT
 ORIGIN

Query Match 75.4%; Score 957; DB 10; Length 2577;
 Best Local Similarity 85.2%; Pred. No. 7.4e-285;
 Matches 1081; Conservative 0; Mismatches 185; Indels 3; Gaps 1;

QY 1 ATGTGGAACACCTGTGTCACAGACAGTGGATTTCTCAGACAGATGGAATGGCTTCTG 60
 DB 112 ATGTGGAACACCTGTGTCACAGACAGTGGATTTCTCAGACAGATGGAATGGCTTCTG 171

IMAGE:4020396, mRNA, complete cds.
 BC031933
 EC031933.1 GI:21594465
 MGC.
 Mus musculus (house mouse)
 Mus musculus
 Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
 Strausberg, R.
 Direct Submission
 Submitted (06-JUN-2002) National Institutes of Health, Mammalian
 Gene Collection (MGC), Cancer Genomics Office, National Cancer
 Institute, 31 Center Drive, Room 11A03, Bethesda, MD 20892-2590,
 USA
 NIH-MGC Project URL: http://mgc.nci.nih.gov
 Contact: MGC help desk
 Email: cgabbs@mail.nih.gov
 Tissue Procurement: Gilbert Smith, Ph.D.
 cDNA Library Preparation: Life Technologies, Inc.
 cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)
 DNA Sequencing by: Baylor College of Medicine Human Genome
 Sequencing Center
 Center code: BCM-HGSC
 Web site: http://www.hgsc.bcm.tmc.edu/cdna/
 Contact: amg@bcm.tmc.edu
 Gunaratne, P.H., Garcia, A.M., Lu, X., Hulyk, S.W., Hale, S.M.,
 Yoon, V.S., Kowis, C.R., Lawrence, S., Martin, R.G., Muzny, D.M.,
 Richards, S., Gibbs, R.A.

Clone distribution: MGC clone distribution information can be found
 through the I.M.A.G.E. Consortium/LLNL at: http://image.llnl.gov
 Series: IRAK Plate: 30 Row: k Column: 3
 This clone was selected for full length sequencing because it
 passed the following selection criteria: Hexamer frequency ORF
 analysis, similarity but not identity to protein.
 Location/Qualifiers
 1. .2577
 /organism="Mus musculus"
 /mol_type="mRNA"
 /strain="CZECH II"
 /db_xref="taxon:10090"
 /clone="MGC:25790 IMAGE:4020396"
 /tissue type="Mammary tumor metastasized to lung."
 MMTV-LTR/Wnt1 model. Expression driven by an MMTV-LTR
 enhancer.
 /clone_lib="NCI CGAP_Lu30"
 /lab_host="DH10B"
 /notes="Vector: pCMV-SPORT6"
 112. .1380
 /codon_start=1
 /product="RIKEN cdna 4632427C23 gene"
 /protein_id="AAH31933.1"
 /db_xref="GI:21594466"
 /db_xref="LocusID:78753"
 /translation="MSEILSRVTVSHRVEIWLILVAYLLORNNSGHLPTKADPE
 AFNVSEIIXKGYPSSEYEVATEDGYILSVNRIPIRGOTRLKKEGSRPVVLQHLGLG
 DASWTSLNPLNSLGFILADAGFDVWNGSRGNTWSRKHKTLISIDQEFWAFSYDEMA
 RFDLPAVINFLQTKQKVVYGVSGQTMGTGFIATFTPELAIHKIMYFALAPIATV
 KYARSCTKELLIPOMKIVLGRQEFLYQTFERCLFTYLCQMLDQICSNIIILL
 GGFNTNNMSRANVYATPAGTSVQNLHWSQAVNSGELRAFQWSETKKQKCNQ
 PTPIRYKRDMMVPTATWTGGQDLSNDDVKTLLSEVNLVYHKNIPFWAHVDFTWG
 LDAPQRYVNIHLKQEPNLPQGTCTRVKL"

QY	61	ATTCTGGTGGCTATATGTTTCAGAGAAATGTAATTCAGTACATATGCCAATTAAGCT	120
Db	172	ATTCTGGTGGCTATATGTTTCAGAGAAATGTAATTCAGTACATATGCCAATTAAGCT	231
QY	121	GTGGAGCCAGAGAACATTCATGATATATAGTGAATATCATCAACATCAAGGCTATCCCTGT	180
Db	232	GGGATCCAGAGAACATTCATGATATATAGTGAATATCATCAACATCAAGGCTATCCCACT	291
QY	181	GAGGAATATGAGTCCCACTGAGTGGTATATCCCTTCTGTTAAACAGGATTCCTGA	240
Db	292	GAGGATATGAGTCCCACTGAGTGGTATATCCCTTCTGTTAAACAGGATTCCTCG	351
QY	241	GGCTAGTGCACCTAAGAGACAGGTTCCAGGCTGTGGTGTACTGAGCATGGCCTA	300
Db	352	GGACAGACAGGTTAAGAGAGGAGATCCAGGCCAGTGGTGTACTGCAACATGGTCTT	411
QY	301	GTTGGAGGCTGAGCACTGGATTTCCAACTGCCCCCAACATAGCCTGGGCTTCACTTG	360
Db	412	CTTGGGAGTCTAGCACTGGATTTCCAACTGCCCCCAACATAGCCTGGGCTTCACTTG	471
QY	361	GCAAGTCTGCTTTTTCAGCTGTGGATGGGAAACAGCAGGGGAAACGCTGGTCTCGAAA	420
Db	472	GCAAGTCTGCTTTTTCAGCTGTGGATGGGAAACAGCAGGGGAAACGCTGGTCTCGAAG	531
QY	421	CACAAGACATCTCCATAGACCAAGATGAGTTCTGGGCTTTCAGTATGATGAGTGGCT	480
Db	532	CACAAGACATCTCCATAGACCAAGATGAGTTCTGGGCTTTCAGTATGATGAGTGGCT	591
QY	481	AGTTTGACCTCTCTGAGTGAATAAATTTTTCAGAAACAGGGCCAGGAAAGATC	540
Db	592	AGTTTGACCTCTCTGAGTGAATAAATTTTTCAGAAACAGGGCCAGGAAAGATC	651
QY	541	TATTATCTGGCTATTCACAGGACCAACATGGGCTTTATGCAATTTTCCACCAATGCA	600
Db	652	TATTATCTGGCTATTCACAGGACCAACATGGGCTTTATGCAATTTTCCACCAATGCA	711
QY	601	GAGCTGCTCAGAAATCAAAATGATTTTTCAGTATGAGTATGAGTATGAGTATGAGT	660
Db	712	GAGCTGCTCAGAAATCAAAATGATTTTTCAGTATGAGTATGAGTATGAGTATGAGT	771
QY	661	GCAAAAGCCCGGACCAATTTTTCAGTATGAGTATGAGTATGAGTATGAGTATGAGT	720
Db	772	GCAAAAGCCCGGACCAATTTTTCAGTATGAGTATGAGTATGAGTATGAGTATGAGT	831
QY	721	GGCAAAAGCAATTTCTGATCAGACCAATTTCTCAGAACTTGTGTTATTTACCTTCT	780
Db	832	GGCAAAAGCAATTTCTGATCAGACCAATTTCTCAGAACTTGTGTTATTTACCTTCT	891
QY	781	GGCCAGGTGATTTCTGATCAGATTTGATGATATATGATTTCTCTGGGTGAGTTCAC	840
Db	892	GGCCAGGTGATTTCTGATCAGATTTGATGATATATGATTTCTCTGGGTGAGTTCAC	951
QY	841	ACCAACATATGACATGAGCGGAGAGTGTATATCTGCTGCCACACTCTTGCTGGAACA	900
Db	952	ACCAACATATGACATGAGCGGAGAGTGTATATCTGCTGCCACACTCTTGCTGGAAC	1011
QY	901	TCTGTGCAAAATATTTCTACACTGGAGCCAGGAGTGAATTTCTGTTAACTTCGGGCA	960
Db	1012	TCTGTGCAAAATATTTCTACACTGGAGCCAGGAGTGAATTTCTGTTAACTTCGGGCA	1071
QY	961	GACTGGGGAGTGAACAACAAATATCTGGAATAATGCAATGCGCACTCTGTAGGTAC	1020
Db	1072	GACTGGGGAGTGAACAACAAATATCTGGAATAATGCAATGCGCACTCTGTAGGTAC	1131
QY	1021	AGAGTTCAGATATGAGCGTCCCTACAGCAATGAGGAGGAGTCAAGGCTGCTTCA	1080
Db	1132	AAAGTTCAGATATGAGCGTCCCTACAGCAATGAGGAGGAGTCAAGGCTGCTTCA	1191
QY	1081	AATCCAGAGAGTGAATAATCTGCTCTCAGGAGTGAACCACTCTACCTACCAAGAT	1140
Db	1192	AATCCAGAGAGTGAATAATCTGCTCTCAGGAGTGAACCACTCTACCTACCAAGAT	1251

QY	1141	ATTCTGAATGGGCTCAGCTGGATTTTCACTCTGGGTTTGGATGCTCTCACCGTATGTAC	1200
Db	1252	ATTCTGAATGGGCTCAGCTGGATTTTCACTCTGGGTTTGGATGCTCTCACCGTATGTAC	1311
QY	1201	AATGAATATCATCCATCTGATGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG	1260
Db	1312	AATGAATATATACATCTGATGAAGC---AGGAGCCCAACCTTCCCAGGGAACCTGCAGG	1368
QY	1261	GCGGTATTG 1269	
Db	1369	GTCAAAATTG 1377	
RESULT 4			
AX521801			
LOCUS			
DEFINITION			
AX521801			
ACCESSION			
AX521801.1			
VERSION			
KEYWORDS			
SOURCE			
ORGANISM			
Homo sapiens (human)			
Homo sapiens			
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;			
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.			
1			
Xiao, Y.			
Human lysosomal acid lipase			
Patent: WO 0236731-A 4 10-MAY-2002;			
Bayer Aktiengesellschaft (DE)			
Location/Qualifiers			
1. 1116			
/organism="Homo sapiens"			
/mol_type="genomic DNA"			
/db_xref="taxon:9606"			
BASE COUNT			
ORIGIN			
Query Match			
Best Local Similarity			
Matches			
Query Match			
Best Local Similarity			
Matches			
Query Match			
Best Local Similarity			
Matches			
Query Match			
Best Local Similarity			
Matches			
Query Match			
Best Local Similarity			
Matches			
Query Match			
Best Local Similarity			
Matches			
Query Match			
Best Local Similarity			
Matches			
Query Match			
Best Local Similarity			
Matches			
Query Match			
Best Local Similarity			
Matches			
Query Match			
Best Local Similarity			
Matches			
Query Match			
Best Local Similarity			
Matches			
Query Match			
Best Local Similarity			
Matches			
Query Match			
Best Local Similarity			
Matches			
Query Match			
Best Local Similarity			
Matches			
Query Match			
Best Local Similarity			
Matches			
Query Match			
Best Local Similarity			
Matches			
Query Match			
Best Local Similarity			
Matches			
Query Match			
Best Local Similarity			
Matches			
Query Match			
Best Local Similarity			
Matches			
Query Match			
Best Local Similarity			
Matches			
Query Match			
Best Local Similarity			
Matches			
Query Match			
Best Local Similarity			
Matches			
Query Match			
Best Local Similarity			
Matches			
Query Match			
Best Local Similarity			
Matches			
Query Match			
Best Local Similarity			
Matches			
Query Match			
Best Local Similarity			
Matches			
Query Match			
Best Local Similarity			
Matches			
Query Match			
Best Local Similarity			
Matches			
Query Match			
Best Local Similarity			
Matches			
Query Match			
Best Local Similarity			
Matches			
Query Match			
Best Local Similarity			
Matches			
Query Match			
Best Local Similarity			
Matches			
Query Match			
Best Local Similarity			
Matches			
Query Match			
Best Local Similarity			
Matches			
Query Match			
Best Local Similarity			
Matches			
Query Match			
Best Local Similarity			
Matches			
Query Match			
Best Local Similarity			
Matches			
Query Match			
Best Local Similarity			
Matches			
Query Match			
Best Local Similarity			
Matches			
Query Match			
Best Local Similarity			
Matches			
Query Match			
Best Local Similarity			
Matches			
Query Match			
Best Local Similarity			
Matches			
Query Match			
Best Local Similarity			
Matches			
Query Match			
Best Local Similarity			
Matches			
Query Match			
Best Local Similarity			
Matches			
Query Match			
Best Local Similarity			
Matches			
Query Match			
Best Local Similarity			
Matches			
Query Match			
Best Local Similarity			
Matches			
Query Match			
Best Local Similarity			
Matches			
Query Match			
Best Local Similarity			
Matches			
Query Match			
Best Local Similarity			
Matches			
Query Match			
Best Local Similarity			
Matches			
Query Match			
Best Local Similarity			
Matches			
Query Match			
Best Local Similarity			
Matches			
Query Match			
Best Local Similarity			
Matches			
Query Match			
Best Local Similarity			
Matches			
Query Match			
Best Local Similarity			
Matches			
Query Match			
Best Local Similarity			
Matches			
Query Match			
Best Local Similarity			
Matches			
Query Match			
Best Local Similarity			
Matches			
Query Match			
Best Local Similarity			
Matches			
Query Match			
Best Local Similarity			
Matches			
Query Match			
Best Local Similarity			
Matches			
Query Match			
Best Local Similarity			
Matches			
Query Match			
Best Local Similarity			
Matches			
Query Match			
Best Local Similarity			
Matches			
Query Match			
Best Local Similarity			
Matches			
Query Match			
Best Local Similarity			
Matches			
Query Match			
Best Local Similarity			
Matches			
Query Match			
Best Local Similarity			
Matches			
Query Match			
Best Local Similarity			
Matches			
Query Match			
Best Local Similarity			
Matches			
Query Match			
Best Local Similarity			
Matches			
Query Match			
Best Local Similarity			
Matches			
Query Match			
Best Local Similarity			
Matches			
Query Match			
Best Local Similarity			
Matches			
Query Match			
Best Local Similarity			
Matches			
Query Match			
Best Local Similarity			
Matches			
Query Match			
Best Local Similarity			
Matches			
Query Match			
Best Local Similarity			
Matches			
Query Match			
Best Local Similarity			
Matches			
Query Match			
Best Local Similarity			
Matches			

Qy	628	TTTGCTTTAGCACCACATAGCCTGTTAAAGCATGCAAAAGAGCCCGGAGCAAAATTTTG	687
Db	553	TTTGCTTTAGCACCACATAGCCTGTTAAAGCATGCAAAAGAGCCCGGAGCAAAATTTTG	612
Qy	688	TTGCTGCCAGATATGATGATCAGGAGTGTGTTGGCAAAAGAAATTTCTGTATCAGACC	747
Db	613	TTGCTGCCAGATATGATGATCAGGAGTGTGTTGGCAAAAGAAATTTCTGTATCAGACC	672
Qy	748	AGATTCTCAGCAAACTGTTTATTACCTTTGTGGCCAGGATTTCTTTGATCAGATTGT	807
Db	673	AGATTCTCAGCAAACTGTTTATTACCTTTGTGGCCAGGATTTCTTTGATCAGATTGT	732
Qy	808	AGTAATATCTGTTACTTCTGGGTGGATCAACACCAACCAATATGAACAT	857
Db	733	AGTAATATCTGTTACTTCTGGGTGGATCAACACCAACCAATATGAACAT	792
Qy	858	-----GAGCCGAGCAAGTGTATATCTGCTGCCACACTCTTGTGGAACATCTGTG	906
Db	793	GTTTGTGTACAGAGCCGAGCAAGTGTATATCTGCTGCCACACTCTTGTGGAACATCTGTG	852
Qy	907	CAAAATATCTTACTGAGCCAG	930
Db	853	CAAAATATCTTACTGAGCCAG	876
RESULT 5			
AK091558			
LOCUS			
DEFINITION			
AK091558			
VERSION			
KEYWORDS			
SOURCE			
ORGANISM			
REFERENCE			
AUTHORS			
TITLE			
JOURNAL			
COMMENT			
FEATURES			
source			

BASE COUNT	847	a	596	c	615	g	924	t
ORIGIN								
Query Match								
Best Local Similarity								
Matches	792							
Conservative	0							
Mismatches	413							
Indels	6							
Gaps	1							
Qy	37	CACAGAAATGGAATGCGCTTCTGATTCTGGTGGCGTATATGTTCCAGAGAAATGTGAAT	96					
Db	672	CACAGAAATGGAATGCGGTCTTGGGGTGGTGGTCTGTTGGTCTCTCTGACCCCTGCAT	731					
Qy	97	TCAG-----TACATATGCCAACTAAAGCTGTGGHACCAGAGCAATTCATGAATATAGT	150					
Db	732	TCAGAGGGGTCTGGAGGGGAAATGACAGCTGTGGATCTGAAACAAACATGAATGTAGT	791					
Qy	151	GAAATCATCAACATCAAGGCTATCCCTGTGAGGAATATGAATCGCACTGAAAGATGGG	210					
Db	792	GAAATTAATCTTACTTGGGGATTCCTAGTGAGGAATACCTAGTTGAGACAGAGATGGA	851					
Qy	211	TATATCCTTTCTGTTAAACAGGATTCCTCGAGCCCTAGTGCAACCTTAAGAACAGAGTTCC	270					
Db	852	TATATCTGTGCTTAAACCGAATTCCTCATGGGAGAGAAACCAATTCCTGACAAAGGTCC	911					
Qy	271	AGSCCTGTGGTGTACTCGAGCATGSCCTAGTTGGAGGTGTAGCAACTGTGATTTCCAAAC	330					
Db	912	AAACAGTTGCTTCTCTGCAACATGGCTTGTGGCAGATCTAGTAACTGGGTCAACAC	971					
Qy	331	CTGCCCAACAATPAGCTGGGCTTCAATCTGGCAGATGCTGGTTTGAAGTGTGGAGTGGG	390					
Db	972	CTTGCAGCAGCAGCGCTGGGCTTCATTTCTGTGATGCTGGTTTGGAGCTGTGGATGGC	1031					
Qy	391	AAACGAGGGGAAACGCTGTCTCGAAACACACAGACACTCTCCATAGACCAAGATCAG	450					
Db	1032	AAACGAGAGGAAATACCTGGTCTCGGAAACATCAGACACTCTAGTTTCTCAGAGTGA	1091					
Qy	451	TTCTGGGCTTTTCAAGTTATGATGAGATGGCTAGTTTGAACCTTCTCGAGTGAATAACT	510					
Db	1092	TTCTGGGCTTTTCAAGTTATGATGAGATGGCAAAATATGACCTACCAAGCTTCCATTAACT	1151					
Qy	511	ATTTTCAGAAACCGGCGCAGGAAAGATCTATTTATGTCGGCTATTCACAGGCGACCAAC	570					
Db	1152	ATTTTCAGAAACCGGCGCAGGAAAGATCTATTTATGTCGGCTATTCACAGGCGACCAAC	1211					
Qy	571	ATGGGCTTTATTTGCAATTTTCCACCATGCGAGCTGGCTCAGAAATCAAAATGTATTTT	630					
Db	1212	ATAGGTTTATAGCATTTTCAAGATCCCTGAGCTGGCTTAAAGGATTTAAATGTTTTT	1271					
Qy	631	GCTTTAGCACCATAGCCACTGTTAAGCTGCAAAAGCCCGGAGCCAAATTTTGTG	690					
Db	1272	GCCCTGGGTCTGTGGCTTCCGTCTGTACTAGCCCTATGGCCCAATTTAGGACGA	1331					
Qy	691	CTGCCAGATATGATGATCAAGGGATTGTTTGGCAAAAGAAATTTCTGTATCAGACCAAG	750					
Db	1332	TTACCAAGATCATCTCATTAAGACATTTTGGAGACAAAGAAATTTCTCCCAAGATGGG	1391					
Qy	751	TTTCTCAGACAACTGTTTATTTACCTTTGTGGCCAGGCTGATTTCTTGATCAGATTTGT	810					
Db	1392	TTTTTGAAGTGGCTGGGTACCCACCGTTTGCATCTCATGTCACTGAAGGAGCTCTGTGGA	1451					
Qy	811	AATATCATCTTACTTCTGGTGGATTTCAACCAACAATATGAACATGAGCCGAGCAAGT	870					
Db	1452	AATCTGTTTTTCTTCTGTGGATTTTAAAGAGAAATTTAAATATGTTAGATGGAT	1511					
Qy	871	GTATATGCTGCCACACACTCTTCTGCAACATCTGTGCAAAATTTCTACCTGGAGCCAG	930					
Db	1512	GTATATACAAACATTTCTCTGCTGGAATCTTGTGCAAAACATGTTTACCTGAGCCAG	1571					
Qy	931	GCAGTGAATCTGGTGAATCTGGGCAATTTGCTGGGAGTGAGACCAAAATCTGGAA	990					
Db	1572	GCTGTTAAATTCAAAAGTTTCAAGCCCTTTGACTGGGAGGAGCAGTGCCAAAGATCAT	1631					
Qy	991	AAATGCAATCAGCCAACTCTCTGTAAGGTACAGAGTACAGATATGACGCTCCCTACAGCA	1050					

Db 1632 CATTACACAGAGTTATCTCCCAATACATGTAAGGACATGCTTTGGCCGACTGCA 1691
 QY 1051 ATGTGGACAGAGGTGAGGCTGCTTCAAAATCCAGAGACGTGAAGATGCTCTCT 1110
 Db 1692 GTCTGAGCGGGGTGACGACTGCTTCCAGATGCTACGAGCTCAATATCTTACTGACT 1751
 QY 1111 GAGGTGACCACTCATCTACCATGAATATCTTGAATGGCTCAGTGGATTCATC 1170
 Db 1752 CAGATCACCACTTGGTGTTCATGAGAGCATTCGGAAATGGAGACATCTTGACTTCATT 1811
 QY 1171 TGGGGTGTGGATGCTCCTCACCGTATGACATGAATCATCCATCTCATGACAGAGAG 1230
 Db 1812 TGGGGCTGTGATGCTTGGAGGCTTTATATAAAATTAATTAATTAATGAGGAATAT 1871
 QY 1231 GAGACCACT 1241
 Db 1872 CAGTGAAGCT 1882

RESULT 6
 AX411234 2481 bp DNA linear PAT 14-JUN-2002

LOCUS
 DEFINITION Sequence 3881 from Patent WO0229103.
 ACCESSION AX411234
 VERSION AX411234.1 GI:21443939

SOURCE Homo sapiens (human)
 ORGANISM
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.

REFERENCE
 1. Alvares, C., Horne, D., Peres-da-Silva, S. and Vockley, J. G.
 Gene expression profiles in liver cancer
 Patent: WO 0229103-A 3881 11-APR-2002;
 GENE LOGIC INC (US)

FEATURES
 source
 1. 2481
 /organism="Homo sapiens"
 /mol_type="genomic DNA"
 /db_xref="taxon:9606"
 /notes="EMBL/GenBank Accession No. 231690"
 BASE COUNT 692 a 477 c 513 g 799 t

Query Match 41.6%; Score 527.8; DB 6; Length 2481;
 Best Local Similarity 66.6%; Pred. No. 1e-151;
 Matches 751; Conservative 0; Mismatches 376; Indels 0; Gaps 0;

QY 115 AAAGCTGTGACCCAGAGCATTCATGAATATTAGTGAATCATCCAAATCAAGGCTAT 174
 Db 112 ACAGCTGTGATCTCTGAAACAAATGATGTGATGAATATCTTACTGGGATTC 171
 QY 175 CCCTGTGAGGAATATGAAGTCGCAACTGAAGATGGGTATATCTTCTGTTAACAGGAT 234
 Db 172 CCTAGTGAAGATACCTAGTTGAGACAGAGATGATATCTGTGCTTAAACCGAAT 231
 QY 235 CCTCGAGCTAGTGCACTTAAGACAGAGTTCCAGGCTGTGTGTTTACTGACGAT 294
 Db 232 CCTCATGGGAGGAAGAACCAATCTGACAAAGGTCCTCAACAGATGCTCTCTGCAAT 291
 QY 295 GGCCTAGTGTGAGGTGTAGCAACTGGAATTCACAACTGCCCAACAAATAGCTGGGCTTC 354
 Db 292 GGCCTGTGGCAGATCTTAGTAATCTGGGTCAACAACTGCCCAACAGAGCTGGGCTTC 351
 QY 355 ATTTGCGAGATGCTGTTTGTAGCTGTGATGGGACAGAGGGGAAACGCTGTCT 414
 Db 352 ATTTGCTGTGATGCTGTTTGTAGCTGTGATGGGCAACAGAGGAATATCTGTCT 411
 QY 415 CGAAACACAGACACTCTCATAGAACCAAGATGAGTCTTGGGCTTTCACTATGATGAG 474
 Db 412 CGAAACATAGACACTCTCAGTTCTCAGGATGAATCTTGGGCTTTCACTATGATGAG 471

QY 475 ATGGCTAGGTTTACCTCTGCAATATAAATTTATTTTGCAGAAACGGCCAGAA 534
 Db 472 ATGGCAATAATGACTTACCAGCTTCCATTAACCTTCTGATATAAATGGCCAGAA 531
 QY 535 AAGATCTATTATGTGGGTATTTCAAGGACCAACATGGGCTTTATTTGCATTTTCCACC 594
 Db 532 CAAGTGATATTATGTGGGTCAATTTCTCAAGGACCACTTAGGTTTATATAGCATTTTCAAG 591
 QY 595 ATGCCAGAGCTGGCTCAGAAAATCAAAATGATTTTGTCTTTAGCAACCATAGCCACTGTT 654
 Db 592 ATCCCTGAGCTGGCTAAAGGATTAAGTTTGTCTTTGCTGGTCTGCTGCTTCCGTC 651
 QY 655 AAGCATGCAAAAGCCCGGACCAAAATTTTGTGCTGCGCAGATATGATGATCAAGGA 714
 Db 652 GCCTTCTGCTAGCCCTATGCCCATAATTTAGGACGATTACCCAGATCATCTCAATAGGAC 711
 QY 715 TTGTTTGGCAAAAAGAAATTTCTGTATCAGACAGATTTCTCAGACAACTTCTTATTTAC 774
 Db 712 TTATTTGGAGACAAAGAAATTTCTCCAGAGTGGCTTTTGAAGTGGCTGGGTACCCAC 771
 QY 775 CTTTGTGGCCAGGTGATTTCTGATCAGATTTGTAGTAATATCATGTTACTTCTGGTGA 834
 Db 772 GTTTGCACTCATGTCTACTGAAGAGCTCTGTGAAATCTCTGTTTCTTCTGTTGGA 831
 QY 835 TTCAACACCAACATATGACATGAGCCGAGCAAGTGTATATGCTGCCCACTCTTGTCT 894
 Db 832 TTTATGAGAGAAATTTTAAATATGTCTAGAGTGGATGATATATACACATTTCTCTCT 891
 QY 895 GGAACATCTCTGCAAAATATTTTACACTGGAGCAGGAGTGAATTTCTGGTGAACCTCCGG 954
 Db 892 GGAATCTCTGCAAAACATGTTTACACTGGAGCAGGCTGTTAAATTTCCAAAGTTTCAA 951
 QY 955 GCATTTGACTGGGGGTGAGACCAAAATCTGAAAATGCAATCAGCCAACTCTCTGTA 1014
 Db 952 GCCTTTGACTGGGGAAGCAGTGCCTCAAGAAATTTTTCATTAACCAAGATTTCTCTCC 1011
 QY 1015 AGGTACAGAGTCAGAGATATGACGCTCCCTACAGCAATGTGGACAGAGGTGAGGCTGG 1074
 Db 1012 ACATACATGTAAGGACATGCTTGTGCCGACTGCTGAGCGGGGTGACGACTGG 1071
 QY 1075 CTTTCAATCCAGACAGGTGAATGCTCTCTGAGGTGAGCACTTCACTCATCTACCAT 1134
 Db 1072 CTTGAGATGCTACGAGCTCAATATCTTACTGACTAGATCACCACCTTGGTGTCCAT 1131
 QY 1135 AAGAAATTTCTGAATGGGCTCAGCTGATTTCAATCTGGGGTTTGGATGCTCTCCCGT 1194
 Db 1132 GAGAGCATTCGGAATGGGAGCATTTGACTTTCATTTGGGGCTGATGCCCTTGGAGG 1191
 QY 1195 ATGTACATGAATCATCTCATCTGATGACAGAGGAGGACCACT 1241
 Db 1192 CTTTATAATAAATTTAATTAATTAATGAGGAATATCAGTGAAGCT 1238

RESULT 7

HLXYACLY
 LOCUS
 DEFINITION H.sapiens (HepG2) LAL mRNA for lysosomal acid lipase.
 ACCESSION Z31690
 VERSION Z31690.1 GI:506430
 KEYWORDS LAL; lipase; lysosomal acid lipase.
 SOURCE Homo sapiens (human)

ORGANISM

Homo sapiens
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.

REFERENCE

1 (bases 1 to 2481)

AUTHORS

Du, H. and Gregory, G.A.

TITLE

Structural Conservation of Putative Functional Motifs between Mouse

JOURNAL

Unpublished

REFERENCE

2 (bases 1 to 2481)

AUTHORS

Du, H.

TITLE

Direct Submission

JOURNAL

Submitted (05-APR-1994) Hong Du, Division of Human Genetics,

Children's Hospital Medical Center, 3333 Bernet Street, Cincinnati,
OH, 45229-3039, USA

FEATURES

Location/Qualifiers
1. .2481
/organism="Homo sapiens"
/mol_type="mRNA"
/strain="HepG2"
/db_xref="taxon:9606"
/clone="clone hLAL"
/cell_type="hepatoma"
/tissue_type="liver"
/clone_lib="human liver HepG2 cDNA library/lambda ZAP"
1. .2481
/gene="LAL"
34. .1233
/gene="LAL"
/citation=[1]
/codon_start=1
/product="lysosomal acid lipase"
/protein_id="CAA83495.1"
/db_xref="GI:506431"
/db_xref="SPTREMBL:Q16529"
/translation="MKMFLGIYVCLVLTLSHSGRGLTAVDPETNMVSEIISVM
GPFSEYLVETDGVILCLNRIPIGRKNSKGPVVPFLOHGLLADSSNWTNLANS
SIGFLADAGDVMNGRNTWSRKHKLTVSVDQFWAPSDYEMAKIDLPAISNPL
NKTGEQYVYVGHSGQTTFGTAFSPQIPELAKRIMFFALGFVASVAFSPMAKGR
LPDLHILKDLFGDKPELPQSAFLKMLGTHVCHVILKELCNLCFLLCGFNERNLNMR
VDVYTHSPAGTSVQNMHLSQAVKQFQAFDMGSSAKNYFHYNQSVPTNYKMDL
VPTAVSGSHDLADVVDVNIILQITNLVPHESIPEWHLDFIWLDPWRLNKII
NLMRKYQ"
34. .96
/gene="LAL"
/citation=[1]
97. .1230
/gene="LAL"
/product="lysosomal acid lipase"
/citation=[1]
79
/gene="LAL"
/citation=[1]
100
/gene="LAL"
/citation=[1]
/replaces="g"
2459. .2464
/gene="LAL"
2480
/gene="LAL"
692 a 477 c 513 g 799 t
BASE COUNT
ORIGIN
Query Match 41.6%; Score 527.8; DB 9; Length 2481;
Best Local Similarity 66.6%; Pred. No. 1e-151;
Matches 751; Conservative 0; Mismatches 376; Indels 0; Gaps 0;
Qy 115 AAAGCTGTGACCCAGAGCATTCATGAATATTAGTGAATATCAATCAAGCTAT 174
Db 112 ACAGCTGTGATCTGAAACAAACATGATGTGATGAATATCTCTTACTGGGATTC 171
Qy 175 CCTGTGAGGAATATGAAGTCGCACTGAAGATGGGTATATCTCTTTTAAACAGGATT 234
Db 172 CCTAGTGAAGATATCTAGTGTGAGACAGAGATGATATATCTGTGCCCTTAAACGGAAT 231
Qy 235 CCTCGAGGCTAGTGACACCTTAAGACAGAGTTCCAGGCTGTGTGTGTACTTGCAGCAT 294
Db 232 CCTATCGGAGGAGAAACCAATCTGACAAAGGTCCCAACACAGTTGTCTTCTGCAACAT 291
Qy 295 GGCTAGTGTGAGGTGTAGCAACTGGATTTCACACCTGCGCCCAACAAATAGCTGGGCTTC 354
Db 292 GGCTTGTGCGAGATTCTAGTAAGTGGGTACAAAACCTTGCACACAGAGCGCTGGGCTTC 351
Qy 355 ATTCTGCGAGATGCTGGTTTGTGCGGTGTGATGGGACAGACAGCGGGAAACGCTGGTCT 414

Db 352 ATTCTGTGTGATGCTGGTTTTCAGCTGTGTGATGGGCAACAGCAGAGGAAATACCTGGTCT 411
Qy 415 CGAAAACACAGACACTCTCCATAGACCAAGATGATGTTCTGGGCTTTCAGTTATCATGAG 474
Db 412 CGGAAACATAAGACACTCTCAGTTTCTCAGGATGAATTTCTGGGCTTTCAGTTATCATGAG 471
Qy 475 ATGGCTAGGTTTTCAGCTTCTCCTCAGTGATAAACTTTTATTTTGCAGAAAACGGCCAGAA 534
Db 472 ATGGCAAAATATGACCTACCACTTCCATTAACTTCTGAATAAACTGGCCAAAGAA 531
Qy 535 AGATCTATATGTGTGGCTATTTCAGGSCACACCATGGGCTTTTATTCATTTCCACC 594
Db 532 CAAGTGTATATGTGGGTCAITTCAGGSCACACCATAGGTTTATAGCATTTTCACAG 591
Qy 595 ATGCCAGAGCTGGCTCAGAAAATCAAAATGTATTTTGTCTTACACCCCATACCTGTT 654
Db 592 ATCCCTGAGCTGGCTAAAGAGATTAATTTTGTCTGCTGGTCTCTGTGGCTTCCGTC 651
Qy 655 AAGCATGCAAAAAGCCCGGACCAAAATTTTGTCTGCCAGATATGATGATCAAGGA 714
Db 652 GCCTTCTGTACTAGCCCTATGGCCAAATTTAGGACGATTAACCATCATCTCATTAGGAC 711
Qy 715 TTCTTTGGCAAAAAGAAATTTCTGTATCAGACCAAGATTTCTCAGACAACTTGTATTAC 774
Db 712 TTATTTGGACAAAGAAATTTCTCCAGAGTGCGTTTTCAGTGGCTGGGTACCCAC 771
Qy 775 CTTTGTGGCCAGGTGATTTCTGATCAGATTTGTAGTAATATCATGTTACTTCTGGGTGA 834
Db 772 GTTTGCACTCATCTCATCTGAGGAGCTCTGTGAAATCTCTGTTTCTTCTGTGTGA 831
Qy 835 TTCAACACCAATATGACATGAGCGGACGAGCAAGTATATGCTGCCACACTCTTGTCT 894
Db 832 TTTAATGAGAGAAATTTAAATATGCTAGTGGATGATATACACATTTCTCTCTGCT 891
Qy 895 GGAACATCTGTGCAAAATATTTCTACACTGAGCAGGAGTGAATTTCTGGTGAATCTCGG 954
Db 892 GGAATTTCTGTGCAAAACATGTTTACACTGGAGCAGGCTGTAAATTTCCAAAAGTTTCAA 951
Qy 955 GCATTTGACTGGGGAGTGAGACCAAAATCTGAAAATGCAATCAGCCAACTCTCTGTA 1014
Db 952 GCCTTTGACTGGGGAGAGTGCACCAAGATTTATTTTATTACCAAGATTTATCTCTCC 1011
Qy 1015 AGGTACAGAGTCAGAGATATGAGGTCCTTACACCAATGTGGAAGAGGTCAGGACTGG 1074
Db 1012 ACATACATGTGAGGACATGCTTGTGCGGACTGCGAGTCTGGAGCGGGGTCACGACTGG 1071
Qy 1075 CTTTCAAAATCCAGAGACGTGAAAATGCTGCTCTCTGAGTGACCAACCTCATCTACCAT 1134
Db 1072 CTTGAGATGCTTACGAGCTCAATATCTTACTGACTCAGATCACCACCTTGGTGTCCAT 1131
Qy 1135 AAGAAATTTCTGAATGGGCTCAGTGGATTTCTATCTGGGGTTGGATGCTCTCACCCT 1194
Db 1132 GAGAGCAATTCGGAATGGGAGCATCTGACTTCTATTGGGCTGGATGCCCTTGGAGG 1191
Qy 1195 ATGTACAAATGAATTCATCTGATGACAGCAGAGGAGACCAACCT 1241
Db 1192 CTTTATAATAAAATTAATTAATTAATGAGGAAATATCAGTGAAGCT 1238
RESULT 8
HSU08464
LOCUS
DEFINITION
ACCESSION
VERSION
KEYWORDS
SOURCE
ORGANISM
REFERENCE
AUTHORS
HSU08464 2481 bp mRNA linear PRI 23-JUN-1994
Human lysosomal acid lipase mRNA, complete cds.
U08464
U08464.1 GI:505052
Homo sapiens (human)
Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
1. (bases 1 to 2481)
Du, H. and Gregory, G.A.

TITLE Structural Conservation of Putative Functional Motifs between Mouse and Human Lysosomal Acid Lipase

REFERENCE Unpublished
2 (bases 1 to 2481)
Du, H.
Direct Submission
Submitted (08-APR-1994) Hong Du, Division of Human Genetics,
Children's Hospital Medical Center, 3333 Burnet Street, Cincinnati,
OH 45229-3039, USA

FEATURES
source Location/Qualifiers
1..2481
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/clone="clone h1AL"
/cell_line="HepG2"
/tissue="liver"
/clone_lib="HepG2 cDNA library/Lambda ZAP"
1..33
34..1233
/codon_start=1
/product="lysosomal acid lipase"
/protein_id="AAB60328.1"
/db_xref="GI:505053"
/translation="MKRFLGLVCLVLTLSHSGRGLTAVDPETNWNSEIIISY
GFPEEYLVEDEGYILCLNRI PHGRNHSDDKPKVFLQHLGLADSNWVNLANS
SLGFIADAGFVDMGNSRGTWSRKHTLSVQDFWAFSDYDEWAKYDLPASINPIL
NKTGQGVYVYVQSGTIGTAFISQIPELAKRIMFFALGPVAVSFCTPMXKIGR
LPHLKLIDFGDEKFLPQSAFLKWLGHVTHVILKELCNLCFLCIGFNERNLNMSR
VDVYTHSPAGTSVQNMHLWSQAVKFKQAPDMGSSAKNYFHYNSQYPTVNVKDL
VPTAVMSGGHDLADYVDVNYLLTQITNLVPHESIPWEHDFEINGLDAPRWLYNKII
NLNRKIQ"

sig_peptide 34..96
mat_peptide 97..1230
variation 79
variation 100
3'UTR 1234..2480
polyA_signal 2459..2464
polyA_site 2480
BASE COUNT 692 a 477 c 513 g 799 t
ORIGIN

Query Match 41.68; Score 527.8; DB 9; Length 2481;
Best Local Similarity 66.68; Pred. No. 1e-151;
Matches 751; Conservatives 0; Mismatches 376; Indels 0; Gaps 0;
QY 115 AAAGCTGTGGACCGAAGCATTCATGAATATTAGTGAATATCCCAACATCAAGGCTAT 174
DB 112 ACAGCTGTGGATCCTGAAACAAACATGAATGTGAGTGAATATCTTACTGGGATTC 171
QY 175 CCCTGTGAGATATGAGTGCACATCGAGATGGGTATATCTTCTGTACAGATT 234
DB 172 CCTAGTGAAGTAACCTAGTGTGACAGAGATGATATCTGTGCCCTTAACCAATT 231
QY 235 CCTGAGGCTTAGTGCACACCTTAAGAACAGGTTCCAGGCTGTGGTGTACTGACGAT 294
DB 232 CCTATGGGAGGAAGAACCTTCTGACAAAGGTCCTCAACAGTTGCTTCTCTGCAACAT 291
QY 295 GGCCTAGTGTGAGGTGTAGCAATCGAATTTCCAACTGCGCCACATATGCTGGCTTC 354
DB 292 GCCTGTGCGGAGATTTCTAGTAATCTGGGTCAAAACCTTGGCAACAGCGCTGGGCTTC 351
QY 355 ATTCTGGCAGATGCTGGTGTGAGTGTGGGAAACAGCAGGGGAAACGCTGTGCT 414
DB 352 ATTCTGTGATGCTGTGTTTGAAGTGTGGATGGGCAACAGCAGGAAATACCTGTCT 411
QY 415 CGAAACACAGACACTCTCCATAGACCAAGATGAGTCTGGGCTTTCAGTTATGATGAG 474
DB 412 CGAAACACAGACACTCTCAGTTTCTCAGGATGAATTTCTGGGCTTTCAGTTATGATGAG 471

QY 475 ATGGCTAGGTTGACCTTCTCTGAGTGATAAACTTTATTTTGAGAAAACGGCCAGGAA 534
DB 472 ATGGCAAAATATGACCTTACAGCTTCCATTTAATTTCTGATTAATACTGGCCAGAA 531
QY 535 AAGATCTATTATGTCGGCTATTCAAGGCGACCAACATGGGCTTTATGCAATTTCCACC 594
DB 532 CAAGTGTATTATGTCGGCTATTCTCAAGGCGACCACTATAGGTTTATAGCATTTTCCACAG 591
QY 595 ATGCCAGAGCTGGCTCAGAAATCAAAATGATTTTGTGCTTTCAGCCATCAGCACTGTT 654
DB 592 ATCCCTGAGCTGGCTAAAGGATTAAGATTTTGTGCTTTCAGCCATCAGCACTGTT 651
QY 655 AAGCATGCAAAAAGCCCGGGGCAAAATTTTGTGCTTTCAGCATATGATGATCAAGGGA 714
DB 652 GCCTTCTGTACTAGCCCTATGCGCAATTAGGACGATTAACAGATCATCTCATTAAGGAC 711
QY 715 TTGTTTGGCAAAAAGGAAATTTCTGATCAGACCAAGATTTCTCAGACAACTTGTATTATAC 774
DB 712 TTAATTTGGAGACAAAGAAATTTCTTCCCAAGATGCGTTTGAAGTGGCTGGGTACCCAC 771
QY 775 CTTTGTGGCCAGGTGATTCTTGTATCAGATTTGTAGTAATAATCATGTTTCTTCTGGGTTGA 834
DB 772 GTTTGCACTCATGTACTAGAGGAGCTCTGTGGAATCTCTGTTTCTTCTGTGTGGA 831
QY 835 TTCAACACCAACATATGACATGAGCGGAGCATGATGATGATGATGATGATGATGATGATGAT 894
DB 832 TTTAATGAGAAATTTAAATATATGCTAGAGTGGATGATATATACACACATTTCTCTGCT 891
QY 895 GGAACATCTGTGCAAAATATTCTACCTGGAGCGAGCAGTGAATTTCTGTGAACTCCGG 954
DB 892 GGAATCTGTGCAAAACATGTTTACCTGGAGCGAGCTGTAAATTTCAAAAGTTTCAA 951
QY 955 GCATTTGACTGGGGAGTGAGACCAAAATCTGAAAAATGCAATCAGCCAACTCTCTGTA 1014
DB 952 GCCTTTGACTGGGAGAGCGATGCGCAAGATTTATTTTCTTACCAACAGAGTTATCTCTCC 1011
QY 1015 AGGTACAGATCAGAGATATGACCGTCCCTACAGCAATGTGGAGAGGTCAGGACTGG 1074
DB 1012 ACATACATGTGAAGGACATGCTTGTGCGGACTGACGTCTGGAGCGGGGTCACGACTGG 1071
QY 1075 CTTTCAATCCAGAGAGCGTGAATGCTGCTCTGAGTGTGACCACTCATCTACCAT 1134
DB 1072 CTTGAGATGCTTACGAGCTCAATATCTTACTGATCAGATCACCACCTTGTGTTCCAT 1131
QY 1135 AAGATATTTCTGAAATGGGCTCAGGTGATTTCTATCTGGGTTTGGATGCTCTCCACCGT 1194
DB 1132 GAGGATTTCCGATGGAGCATCTTGACTTTTATTTGGGCTGGATGCCCTTGGAGG 1191
QY 1195 ATGTACATGAATCATCTCATCTGATGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 1241
DB 1192 CTTTATAATAAATTTATTAATCTAATGAGGAATATCAGTGAAAGCT 1238

RESULT 9
HSLAL HSLAL 2626 bp mRNA linear PRI 25-FEB-1994
LOCUS H.sapiens mRNA for lysosomal acid lipase.
DEFINITION X76488
ACCESSION X76488
VERSION GI:434305
KEYWORDS lysosomal acid lipase.
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE 1
AUTHORS Ameis, D., Merkel, M., Eckerskorn, C. and Gretchen, H.
TITLE Purification, characterization and molecular cloning of human
hepatic lysosomal acid lipase
JOURNAL Eur. J. Biochem. 219 (3), 905-914 (1994)
MEDLINE 94155897
PUBMED 8112342
REFERENCE 2 (bases 1 to 2626)

475	ATGGCTAGGTTTGACCTTCTGCGAGTGATFAAACCTTTATTTTGGCAGAAAACGGCCACGAA	534
584	ATGGCAAAATATGACCTACCAGCTTCATTAACTTCATTCTGTAATAAACTGGCCAAAGAA	643
535	AAGATCTATTATATGTCGGCTATTACAGGGCCACACCATGGGCTTTATTGTCATTTTCCACC	594
644	CAAGTGATATATGTTGGGTCACTTCTCAAGGCACCACTATAGTATTATAGCATTTTTCACAG	703
595	ATGCCAGAGCTGGCTCAGAAAAATCAAAATGTATTGCTTTTAGCACCCATAGCCACTGTT	654
704	ATCCCTGAGCTGGCTAAAGAGATTAAATGTTTTTTTGGCCCTGGGCTCTGTTGGCTTCCGTC	763
655	AAGCATGCAAAAAGCCCGGACCCAAATTTTGTGCTGCGACATATGATGATCAAGGGA	714
764	GCCTTCTGTACTAGCCCTATGGCCAAATTAGGAGGATTACAGATCATCTCATTAAGGAC	823
715	TGTTTGGCAAAAAGAAATTTCTGTATCAGACAGAGATTTCTCAGACAACATGTTGTTATTTAC	774
824	TATTTGGAGACAAAGAAATTTCTTCCCGCAGAGTCGGTTTTTGAAGTGGCTGGGTACCCAC	883
775	CTTTGTGGCCAGGTGATTCTTGATCAGATTGTGTAATATCATGTTACTTCTGGGTGGA	834
884	GTTTGCACTCATGTCATACCTGAAGAGCTCTGTGGAAAATCTCTGTTTCTTCTGTGTGGA	943
835	TTCAACACCAACAATATGAACATGAGCGGACGAAGTGTTATATCTGTCGCCACACTCTTGCT	894
944	TTTAATGAGAGAAATTTAAATATGCTTAGAGTGGATGATATATACAACACATTTCTCTGCT	1003
895	GGAACTCTGTGCAAAATATTCTACACTGGAGCCAGGCACTGAATTTCTGTTGAATCTCCG	954
1004	GGAACTCTGTGCAAAAATGTTTACACTGGAGCCAGGCTGTTAAATTTCCAAAAGTTTCAA	1063
955	GCATTTGACTGGGAGTGAGACCAAAAATCTCGMAAAATGCAATCAGCCAACTCTCTGTA	1014
1064	GCCTTTGACTGGGAAGCAGTGCCCAAGATTTATTTTCAATACACAGAGTTATCTCTCC	1123
1015	AGGTACAGAGTCAGAGATATGACGGTCCCTACAGCAATGTGGACAGAGGTACAGACTGG	1074
1124	ACATACAATGTGAAGGACATGCTTTGTGCCAGCTGCACTCTGGAGCGGGGTACGACTGG	1183
1075	CTTTCAAATCCAGAGAGCGTGAAAATGCTGCTCTCTGAGGTGACCAACCTCATCTACCAT	1134
1184	CTTGCAATGTCTACGACGTCATATCTTACTGACTCAGATCACCACACTTGGTGTTCAT	1243
1135	AAGAAATTTCTGGAATGGGCTCAGTGGATTTCAATCTGGGTTTGGATGCTCTCTACCGT	1194
1244	GAGAGCAATCCGGAATGGAGCATCTTGAATTTTCAATTTGGGCGCTGGATGCCCTTGGAGG	1303
1195	ATGTCATATGAATCATCCATCTGATGACGAGGAGGAGACCACT	1241
1304	CTTTATAATAAATTTAATCTAATGAGGAATATCAGTGAAGCT	1350
RESULT 10		
BD094072		
LOCUS	BD094072	
DEFINITION	Shear stress-responsive DNAs.	
ACCESSION	BD094072	
VERSION	BD094072.1	GI:22639660
KEYWORDS	WO 0125427-A/33.	
SOURCE	Homo sapiens (human)	
ORGANISM	Homo sapiens	
REFERENCE	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo. 1 (bases 1 to 2493)	
AUTHORS	Nojima H., Yoshisue H., Obayashi M., Ota T., Kawabata A., Sakurada K., Kuga T., Sekine S., Nakamura Y. and Sugan S.	
TITLE	Shear stress-responsive DNAs	
JOURNAL	Patent: WO 0125427-A 33 12-APR-2001. KYOWA HAKO KOGYO CO LTD, HIROSHI NOJIMA, HAJIME YOSHISUE, MASAYA OYASHI, TOSHIO OTA, AYAKO KAWABATA, KAZUHIRO SAKURADA, TETSURO KUGA, SUSUMU SEKINE, YUSUKE NAKAMURA, SUMIO SUGANO OS Homo sapiens (human)	
COMMENT		

PN WO 0125427-A/33
 PD 12-APR-2001
 PF 02-OCT-2000 WO 2000JP006840
 PR 01-OCT-1999 JP 99P 280976
 PI HIROSHI NOJIMA, HAJIME YOSHISUE, MASAYA OBAVASHI, TOSHIO OTA, PI
 AYAKO KAWABATA,
 PI KAZUHIRO SAKURADA, TETSURO KUGA, SUSUMU SEKINE, YUSUKE NAKAMURA,
 PI SUMIO SUGANO
 PC C12N15/12, C07K14/435, C07K16/18, C12P21/02, C12Q1/68, A61K38/00,
 PC A61K39/395,
 PC A61K48/00, A61P9/10, G01N33/50, G01N33/53
 CC
 FT
 Key Location/Qualifiers
 CDS Location/Qualifiers
 1. .2493
 /organism="Homo sapiens"
 /mol_type="genomic DNA"
 /db_xref="taxon:9606"
 692 a 484 c 516 g 801 t

Query Match 41.5%; Score 526.2; DB 6; Length 2493;
 Best Local Similarity 66.8%; Pred. No. 3.1e-151;
 Matches 751; Conservative 0; Mismatches 376; Indels 0; Gaps 0;

QY 115 AAGCTGTGACCCAGAGCAATTCATGATATTAAGTGAATATCAATCAATCAAGGCTAT 174
 DB 119 ACAGCTGTGGATCTCTGAAACAAACATGAATGTGAGTGAATATCTCTTACTGGGATTC 178

QY 175 CCCTGTGAGCAATGAAGTCGCACTGAAGATGGGTATATCTCTGTAAACAGGATT 234
 DB 179 CCTAGTGAGGAATACCTAGTTGAGACAGAGATGATATATCTGTGCTTAAACGAAT 238

QY 235 CCTCGAGGCTAGTGCACCTTAAGACAGAGTTCAGGCTGTGTGTATCTGACGAT 294
 DB 239 CCTCATGGGAGGAAGAACCTTCTGACAAAGTCCCAACACAGTGTCTTCTCTGCAACAT 298

QY 295 GGCTAGTGTGGAGTGTACCACTGATTTCCAACTGCTCCCAACCAATAGCTGGGCTTC 354
 DB 299 GGCTGTGCGAGATTCAGTAACTGGTCAAAACCTTCCCAACAGAGCTGGGCTTC 358

QY 355 ATTCGCGAGATGCTGTTTTGACGTTGATGGGAGCAAGAGGGGAAACGCTGGTCT 414
 DB 359 ATTCCTGCTGATGCTGTTTTGACGTTGATGGGAGCAAGAGAGGAAATACCTGGTCT 418

QY 415 CGAAACACACAGACACTCTCATAGACCAAGATGAGTCTGGGCTTTCAGTTATGATGAG 474
 DB 419 CGGAAACATAGACACTCTCAGTTTCTCAGGATGAATCTGGGCTTTCAGTTATGATGAG 478

QY 475 ATGGCTAGGTTTGACCTTCTGCGAGTGAATAAATTTATTTGACAGAAACGGGCCAGGAA 534
 DB 479 ATGGCAAAATATGACCTACAGCTTCCATTAATCTTCAATCTGAATAAACTGGCCAAAGAA 538

QY 535 AAGATCTATATGTCGGCTATTCACAGGCAACACAGTGGCTTTATGCAATTTCCACC 594
 DB 539 CAAGTGATATATGTCGGCTATTCACAGGCAACACAGTGGCTTTATGCAATTTCCACC 598

QY 595 ATGCCAGAGTGGCTCAGAAATCAAAATGATTTTGTCTTATGACCCATAGCACTGTT 654
 DB 599 ATCCCTGAGCTGGCTTAAAGATTAATATGTTTTTGGCCCTGGGCTCTGCTCCGTC 658

QY 655 AAGCATGCANAAAGCCCGGACCAAAATTTTGTGTGTCGAGATATGATGATCAAGGGA 714
 DB 659 GCCTTCTGACTAGCCCTATGGCCAAATTTAGAGGATTAACAGATCATCTCATTAAGGAC 718

QY 715 TTGTTTGGCAAAAGATTTCTGTATCAGACAGATTTCTCAGACAACTGTTATTTAC 774
 DB 719 TTATTTGAGACAAAGATTTCTTCCCAAGAGTGGTTTTGAGTGGCTGGTACCCAC 778

QY 775 CTTTGTGGCCAGGTCATCTTGTATCAGATTTGTATGATATCAATGTTACTTCTGGTGA 834
 DB 779 GTTTGCACTCATGTACTAGAGGAGCTCTGTGGAATCTCTGTTTCTTCTGTGTGGA 838

QY 835 TTCAACACCAACAATATGAACATGAGCCGAGCAAGTGTATATGCTGCCACACTCTTGTCT 894
 DB 839 TTTAATGAGAGAAATTTAATATATGCTAGAGTGTATATATACACACATCTCTCTGCT 898

QY 895 GGAACATCTGTGCAAAATATTTACACTGGAGCCAGGAGTGAATTTCTGGTGAACCTGG 954
 DB 899 GGAACATCTGTGCAAAACATGTTACACTGGAGCCAGGCTGTTAAATTTCCAAAGTTTCAA 958

QY 955 GCATTTGACGTGGGAGTGAACCAAAATCTGAAAAATGCAATCAGACCAACTCTCTGTA 1014
 DB 959 GCCTTTGACGTGGGAGGAGTGCACAGAAATTTATTTCAATACACACAGAGTTATCTCTCC 1018

QY 1015 AGGTACAGATGACAGATATGACGGTCCCTACAGCAATGTGGACAGAGTCAAGACTGG 1074
 DB 1019 ACATACAAATGTGAAGGACATGCTTGTCCGACTGCACTGTGGAGCGGGGTACAGACTGG 1078

QY 1075 CTTTCAATCCAGAGACGTGAAAAATGCTCTCTGAGGTGACCAACCTCATCTACCAT 1134
 DB 1079 CTTGCAATGCTACGACGTCATATCTTACTGACTCAGATCACCACCTTGGTGTCCAT 1138

QY 1135 AAGATATCTCTGAATGGGCTCAGTGGATTTCACTGGGGTTTGGATGCTCTCACCCT 1194
 DB 1139 GAGAGCATTCGGAAATGGGAGCATCTTGAATTTCAATTTGGGGCTGGATGCCCTTGGAG 1198

QY 1195 ATGTACATGAATCATCCATCTGATGACGACGAGGAGACCAACCT 1241
 DB 1199 CTTTATATAAATTTAATCTAATGAGGAATATCAGTGAAGCT 1245

RESULT 11
 HUMLIPCHL
 LOCUS Human lysosomal acid lipase/cholesterol esterase mRNA, complete cds.
 DEFINITION
 M74775
 M74775.1 GI:187151
 lysosomal acid lipase/cholesterol esterase.
 KEYWORDS
 SOURCE Homo sapiens (human)
 ORGANISM Homo sapiens
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
 REFERENCE
 1 (Bases 1 to 2493)
 Anderson, R.A. and Sando, G.N.
 Cloning and expression of cDNA encoding human lysosomal acid
 lipase/cholesterol ester hydrolase. Similarities to gastric and
 lingual lipases
 J. Biol. Chem. 266 (33), 22479-22484 (1991)
 J204232
 MEDLINE
 PUBMED
 COMMENT Original
 FEATURES
 source
 1. .2493
 /organism="Homo sapiens"
 /mol_type="mRNA"
 /db_xref="taxon:9606"
 /map="10"
 /tissue_type="fibroblast"
 1. .2493
 /gene="LIPA"
 /note="G00-120-153"
 41. .1240
 /gene="LIPA"
 /EC number="3.1.1.13"
 /note="alternate initiator ATG may be at position 47"
 /codon_start=1
 /product="lysosomal acid lipase/cholesterol esterase"
 /protein_id="AA59519.1"
 /db_xref="GI:187152"
 /db_xref="GDB:G00-120-153"
 /translation="MKRFLGLVCLVLPVHSEGGGKLTAVDPETNMNVSELIISW"

QY	775	CTTTGTGCCAGGTGATCTTGTATGACGATTGTAGTATATATCATGCTTACTTCTGGGTGGA	834
Db	779	GTTTGCACTCATGTCATCATGGAAGGAGCTCTGTGGAATCTCTGTCTTCTGTGTGGA	838
QY	835	TTCAACACCAACATATGAACATGAGCCGACGAAGTGTATATGTCGCCACACACTCTTGCT	894
Db	839	TTTAATGAGAGAAATTTAAATATGTCTGAGTGGATGTATATACAAACATTTCTCTGCT	898
QY	895	GGACATCTGTGCAAAATATTTCTACACTGGAGCCAGGAGTGAATCTGGTGAACCTCGG	954
Db	899	GGAACTTCTGTGCAAAACATGTTTACACTGGAGCCAGGCTGTTAAATTCCTAAAGTTTCAA	958
QY	955	GCATTTGACTGGGGGAGTGGAGACCAAAAATTCGAAAAATGCAATCAGCAACCTCTCTGA	1014
Db	959	GCCTTTGACTGGGGAAGCAGTCCCAAGAAATTTATTTTCAATACACAGAGTTATCTCTCC	1018
QY	1015	AGGTACAGATCAGAGATATGACGGTCCCTACAGCAATGTGACAGGAGGTCAGACTGG	1074
Db	1019	ACATACAAATGTGAAGGACATGCTTGTGGCCAGTCTGAGCGGGGGTCAGACTGG	1078
QY	1075	CTTTCAAATCCAGAAAGACGTGAAAAATGCTCTCTGAGGTACCAAACTCATCTACCAT	1134
Db	1079	CTTGCAATGCTACGACGTCAATATCTTACTGACTCAGATCACCACTTGGTGTCCAT	1138
QY	1135	AGAAATATTTCTGAATGGGCTCAGTGGATTTTCACTGGGGTTTGGATGCTCTCACCGT	1194
Db	1139	GAGAGCATTCGGAAATGGGAGCATCTTGACTTCTTCTTGGGGCTGGATGCCCTTTGGAGG	1198
QY	1195	ATGTACATGAAATCATCTCATCTGATGACGAGGAGACCAACT	1241
Db	1199	CTTTATATATAATTTATATCTATGNGGAAATATCAGTGAAGCT	1245
RESULT 12			
BC012287			
LOCUS			
DEFINITION			
ACCESSION			
VERSION			
KEYWORDS			
SOURCE			
ORGANISM			
REFERENCE			
AUTHORS			
TITLE			
JOURNAL			
REMARK			
COMMENT			

passed the following selection criteria: matched mRNA gi: 434305.

FEATURES
source

Location/Qualifiers
1..2586
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/clones="MGC:5229 IMAGE:2900168"
/tissue_type="placenta, choriocarcinoma"
/clone_id="NH MGC_10"
/lab_host="DH10B"
/note="Vector: pCMV-SPORT6"
101..1300
/codon_start=1
/product="Similar to lipase A, lysosomal acid, cholesterol
esterase (Wolman disease)"
/protein_id="BAH2287.1"
/db_xref="GI:1512672"
/translation="MKRFLGLVCLVLPVHSEGGGKLTALDPETNMNVSIIISYW
GFPSEYLVETEDYIILNRIHPRKXHSQDPKPVVFLQHLGLADSNWNTNLANS
SLGFTLADAGDVHMSGRNWTSPKXKHTLSVQDEFWAFSDYEMAKYDLPASINFL
NKTGOBYVYVGHSGQTTIGFAPSOIPELAKRIKMPFALGPVAVAFCTSPMKLGR
LPDLIKDLKDFELPQSAFLKWLGHVTHVILKELCNLCLFLGFEVNLNMR
VDVYTHSPAGTSVQVHWSQAVKFOAFDWSAKNYPHYNOSYPTVNVKML
VPTAVSGHDWLDADVTDVNIILQIINLVFHSIPEWHLDFINGLDAPWRLNKLII
NMRRYQ"

CDS

736 a 511 c 531 g 808 t
Query Match 41.3%; Score 524.6; DB 9; Length 2586;
Best Local Similarity 66.5%; Pred. No. 9.9e-151;
Matches 750; Conservative 0; Mismatches 377; Indels 0; Gaps 0;

BASE COUNT
ORIGIN

115 AAAGCTGTGGACCCAGAACGATTCATGAATATTAGTGAATCATCCAAACATCAAGGCTAT 174
179 ACAGCTTTGGATCTCTGAAACAAACATGAATGTGAGTGAATATCTCTTACTGGGATTC 238
175 CCCTGTGAGGATATGAGTGCACACTGAGTGGTATATATCTCTTCTGTAAACAGATT 234
239 CCTAGTGGAGTAATCCTAGTTGAGACAGAGATGATATATCTGTGCTTAAACGAAAT 298
235 CCTCAGGCGCTAGTGCACACCTAAGAAAGACAGGTTCCAGGCTGTGGTGTACTGACGAT 294
299 CCTCATGGGAGGAACCACTCTGACAAAGGTCCAAACCAAGTTGTCTCTCTGCAACAT 358
295 GGCCTAGTTGAGGTGCTAGCACTGATGATTTCCAACTGCGCCCAACATAGCTGGGCTTC 354
359 GGCCTGCTGGCAGATCTTAGTAACCTGGGTCAAAACCTTTGCAACAGACAGCTGGGCTTC 418
355 ATTCTGGCAGATGCTGGTGTGACGTGTGGATGGGAAACAGCAGGGAACACCTGGTGTCT 414
419 ATTCTGTGATGCTGTTTGGCTGTGGATGGGCAACAGCAGAGGAATACCTGGTCT 478
415 CGAAACACAGACACTCTCCATAGACCAAGATGATGTTCTGGGCTTTCAGTTATGATGAG 474
479 CGAAACACATAGACACTCTCAGTTTCTCAGGATGAATTTCTGGGCTTTCAGTTATGATGAG 538
475 ATGGCTAGGTTTGGACCTTCTCAGTGTAACTTTATTTTTCAGAAAACGGGCGAGAA 534
539 ATGGCAAAATATGACTACAGCTTCCATTAATCTTCACTTGAATAAACCTGGCCAGAA 598
535 AAGATCTATATGCTGGCTATTCACAGGCGCAACCATGGGCTTTATTCATTTTCCACC 594
599 CAAGTGATATATGCTGGCTATTCACAGGCGCAACCATAGGTTTATAGCATTTTTCACAG 658
595 ATGCCAGCTGGCTCAGAAATCAAAATGATATTTTGTGTAGCAACCATAGCCACTGTT 654
659 ATCCCTGAGCTGGCTAAAGGATTAATATGTTTTTGGCCCTGGGTCTGGCTTCCGTC 718
655 AAGCATGCAAAAGCCCGGAGCCAAATTTTGTGCTGCCAGATATGATGATCAAGGGA 714
719 GCCTTCTGACTAGCCCTATGCCAAATTAGGACGATTACAGATCATCTCATTAAGGAC 778
715 TTGTTGGCAAAAGAAATTTCTGATATCAGACCAAGATTTCTCAGCAACACTGTTATTAC 774

Db 779 TTAATTTGAGACAAAGAAATTTCTTCCCAGAGTGGCTTTTGAAGTGGCTGGGTACCCAC 838
QY 775 CTTTGTGCCAGGTGATCTTGTATCAGATTGTGTAGTAATATATCATGTTACTCTGGGTGA 834
Db 839 GTTTCACATCATGTCATCTACTGAAGAGCTCTGTGGAATCTCTCTTTCTCTGTGGGA 898
QY 835 TTCAACACCAACAATATGAACATGAGCGGAGCAAGTGTATATGTCGCCACACATCTTGTCT 894
Db 899 TTTAATGAGAGAAATTTAAATATGTCTAGAGTGTATATATCAACACATTTCTCTGCT 958
QY 895 GGAACATCTGTGCAAAATATTCTACATGAGGACGAGCAGTGAATTTCTGGTGAATCTCGG 954
Db 959 GGAATCTCTGTGCAAAACATGTTACCTGGAGCCAGGCTGTAAATTTCCAAAAGTTTCAA 1018
QY 955 GCATTTGACTGGGGAGTGAGACCAAAAATCTGGAATAATCAATCAGCCAACTCTCTGTA 1014
Db 1019 GCCTTTGACTGGGAGGAGCAGTGCCAAAGTAATTTTCTATTACACAGAGTATCTCTCC 1078
QY 1015 AGGTACAGAGTCAGAGATATGACGGTCCCTACAGCAATGTGGACAGGAGTCAGAGCTGG 1074
Db 1079 ACATACAATGTGAAGGACATGCTTGTGCCGACTGCACTGTGGAGCGGGGTCAAGACTGG 1138
QY 1075 CTTTCAATCCAGAGAGCTGAAATGCTGCTCTGAGGTGACCAACCTCATCTACCAT 1134
Db 1139 CTTGAGATGTTACAGCTCAATATCTTACTGACTCAGATCAACACTTGTGTTCAT 1198
QY 1135 AAGAATATTCTGAATGGGCTCACTGGATTTCACTGGGGTTTGGATGCTCTCCACCGT 1194
Db 1199 GAGAGCATTCGGATGGGAGCATCTTGACTTTCATTTGGGGCTGGATGCCCTTTGGAGG 1258
QY 1195 ATGTACAATGAATCATCCATCTGATGACGAGGAGGAGACCAACCT 1241
Db 1259 CTTTATAATAAATTAATTAATCTAATGAGGAATATCAGTGAAGCT 1305

RESULT 13

LOCUS A26689 1378 bp DNA linear PAT 05-APR-1995
DEFINITION Precursor of rabbit gastric lipase coding sequence.
ACCESSION A26689
VERSION A26689.1 GI:905029
KEYWORDS synthetic construct
SOURCE synthetic construct
ORGANISM artificial sequences.
REFERENCE 1 (bases 1 to 1378)
AUTHORS Benicourt, C., Blanchard, C. and Junien, J.L.
TITLE Recombinant gastric lipase from rabbit and pharmaceutical
compositions
JOURNAL Patent: EP 0542629-A 9 19-MAY-1993;
INSTITUT DE RECHERCHE JOUVEINAL
FEATURES
source
1..1378
/organism="synthetic construct"
/mol_type="genomic DNA"
/db_xref="taxon:32630"
BASE COUNT 397 a 287 c 276 g 418 t
ORIGIN

Query Match 38.6%; Score 489.2; DB 6; Length 1378;
Best Local Similarity 65.0%; Pred. No. 8.6e-140;
Matches 722; Conservative 0; Mismatches 388; Indels 0; Gaps 0;
QY 125 ACCCAGAGCATTCATGAATATTAGTGAATCATCCAAATCAAGGCTATCCCTGTGAGG 184
Db 136 ACCCTGAAGTGAATATGAATATTAGTCAATGATTTCTTCTGAGGATACCAAGTGA 195
QY 185 AATATGAGTTCGCACTGAAGATGGGTATATCTTCTGTGTTAAAGGATTTCTCCAGGCC 244
Db 196 AATATGAGTTCGCACTGAAGATGGGTATATCTTCTGAGGTAAACAGAAATTTCTTATGGA 255
QY 245 TAGTGCAACCTAAGAAGACAGAGTTCAGGCTGTGGTGTACTGACGATGCGCTAGTTG 304

256 AGAAAAATTCAGGAAACAGAGCCAGAGACCGTGTATTTTTCAGCATGGTTGCTTG 315
 305 GAGGTCTAGCACTGGATTTCCAACTGCGCCCAAAATAGCTGGCTTCATTCGGCAG 364
 316 CATGACCTCAAACTGGATCTCAACCTGCGCCCAAAATAGCTGGCTTCATTCGGCAG 375
 365 ATGCTGTTTTCAGCTGTGGATGGGAAACAGAGGAGAAACCCCTGGCTTCGAAACACA 424
 376 ATGCTGTTTTCAGCTGTGGATGGGAAACAGAGGAGAAACCCCTGGCTTCGAAACACA 435
 425 AGACACTCTCCATAGCAAGATAGTTCCTGGCTTCAGTATGATGAGATGGTAGGT 484
 436 TATATATTCACAGACTCCGTTGAATTCCTGGCTTCAGTATGATGAGATGGTAGGT 495
 485 TTGACCTTCCTGAGTGAATAAATTTTATTTTCAGAAACCGGCGAGGAAAGATCTATT 544
 496 ATGACCTTCAGCCCAATGACTTCATTGTAAGGAACTGGACAGGAAAGCTGCACT 555
 545 ATGTCGGCTATTCACAGGAGCACCACATGGCTTTTATTCATTTTCCACCATGCCAGAGC 604
 556 ATGTTGGTCAATTCACAGGAGCACCACATGGCTTTTATTCATTTTCTCAAAATCCCAAGC 615
 605 TGCTCAGAAATCAAAATGATTTTTCCTTTCAGCAACCCATAGCTGTTTAAAGATGCAA 664
 616 TGCTGAAAGATCAAAATGATTTTTCCTTTCAGCAACCCATAGCTGTTTAAAGATGCAA 675
 665 AAGCCCGGAGCAAAATTTTTCCTGCGCAGATATGATGATCAAGGATGTTTGGCA 724
 676 AAGCCCTGTGAAACAAATTTTTCCTTTCAGCAACCCATAGCTGTTTAAAGATGCAA 735
 725 AAAAAAATTTCTGATCAGACAGATTTCTCAGCAACCTGTTTATTCCTTTCGGCC 784
 736 AAAAAAATTTCTGATCAGACAGATTTCTCAGCAACCTGTTTATTCCTTTCGGCC 795
 785 AGTGATTTCTGATCAGATTTTTCAGCAACCTGTTTATTCCTTTCGGCC 844
 796 GTGAGACTGTAATGTAATTTTCAGCAACCTGTTTATTCCTTTCGGCC 855
 845 ACAAATGAACTGAGCCAGCAAGTGTATGCTGCCACACTCTTCTGTAAGATCTG 904
 856 CAAACTGAACTGAGCTGCTGAGTGTATGCTGCCACACTCTTCTGTAAGATCTG 915
 905 TGCAAAATTTCTGATCAGATTTTTCAGCAACCTGTTTATTCCTTTCGGCC 964
 916 TTCAAAATTTCTGATCAGATTTTTCAGCAACCTGTTTATTCCTTTCGGCC 975
 965 GGGGAGTGAGACCAAAATTTTCAGCAACCTGTTTATTCCTTTCGGCC 1024
 976 GGGGAGTGAGACCAAAATTTTCAGCAACCTGTTTATTCCTTTCGGCC 1035
 1025 TCAGAGATATGAGCTGCTTCAGCAACCTGTTTATTCCTTTCGGCC 1084
 1036 TGACCCCATGATGTCGCAATTTTCAGCAACCTGTTTATTCCTTTCGGCC 1095
 1085 CAGAGACGTGAAATGCTGCTTCAGCAACCTGTTTATTCCTTTCGGCC 1144
 1096 CCAAGATGAGCTGCTGCTTCAGCAACCTGTTTATTCCTTTCGGCC 1155
 1145 CTGAATGGGCTCAGCTGATTTTCAGCAACCTGTTTATTCCTTTCGGCC 1204
 1156 TTCATATATCACTGATTTTTCAGCAACCTGTTTATTCCTTTCGGCC 1215
 1205 AAATCATCCATCTGATGAGCAGGAGGA 1234
 1216 AAATTTTCTATGATGCGAAGATATAA 1245

RESULT 14
 A26690/c
 LOCUS
 DEFINITION
 Precursor of rabbit gastric lipase (comp.).
 A26690
 A26690

A26690.1 GI:905030
 synthetic construct
 synthetic construct
 artificial sequences.
 1 (bases 1 to 1378)
 Benicourt, C., Blanchard, C. and Junien, J.L.
 Recombinant gastric lipase from rabbit and pharmaceutical
 compositions
 Patent: EP 0542629-A 10 19-MAY-1993;
 INSTITUT DE RECHERCHE JOUVEINAL
 Location/Qualifiers
 1. 1378
 /organism="synthetic construct"
 /mol_type="genomic DNA"
 /db_xref="taxon:32630"
 BASE COUNT 418 a 276 c 287 g 397 t
 ORIGIN
 Query Match 38.6%; Score 489.2; DB 6; Length 1378;
 Best Local Similarity 65.0%; Pred. No. 8.6e-140;
 Matches 722; Conservative 0; Mismatches 388; Indels 0; Gaps 0;
 QY 125 ACCCAGAGCATTCATGAATATTAGTGAATATCAATCAATCAAGCTATCCCTGTGAGG 184
 DB 1243 ACCCTGAAGTGAATATGAATATTAGTCAGATGATTTCTACTGGGATACCCAAAGTGAA 1184
 QY 185 AATATGAAGTGGCAACTGAAGATGGGTATATCTTTCTGTAAACAGATTTCTCGAGGCC 244
 DB 1183 AATATGAAGTGGTGAAGATGGTATATCTTCTGAGGTTAAACAGATTTCTTATGGGA 1124
 QY 245 TAGTGCACTTAAGACAGAGGTTCCAGGCTGTGTGTACTGACGATGGCTAGTTG 304
 DB 1123 AGAAAAATTCAGGAAACAGAGGCCAGAGACCGCTGTATTTTTCAGCAGTGGTTGCTTG 1064
 QY 305 GAGTGCTAGCAACTGGAATTTCCAACTGCTCCAACTAGCTGGCTTCATTTCTGGCAG 364
 DB 1063 CATCAGCCTCAAACTGATCTCCAACTGCTCCAACTAGCTGGCTTCATTTCTGGCAG 1004
 QY 365 ATGCTGTTTTCAGCTGTGGATGGGAAACAGAGGAGAAACCGCTGCTCGAAACACA 424
 DB 1003 ATGCTGTTTTCAGCTGTGGATGGGAAACAGAGGAGAAATACCTGCTCAGAAAGAACT 944
 QY 425 AGACACTCTCCATAGACCAAGATGAGTTCTGGCTTTTCACTTATGATGAGATGGCTAGGT 484
 DB 943 TATATATTACCCAGACTCCGTTGAATTTCTGGCTTTTATGATGAAATGGCTAAT 884
 QY 485 TTGACCTTCTGAGTGAATAAATTTTATTTTTCAGAAACCGGCGCAGGAAAGATCTATT 544
 DB 883 ATGACCTTCCAGCCCAAAATTTGACTTTCATTTGAAAGGAACTGGACAGGAAAGCTGACT 824
 QY 545 ATGTCGGCTATTTCAGAGGCAACCACTGCTTTTATTTGCAATTTTCCACATCCAGAGC 604
 DB 823 ATGTTGTCATTTCCAGGCGCACCACTTATGCAATTTTATGCTTTTCTACAAATCCAGC 764
 QY 605 TGGCTCAGAAATCAAAATGATTTTTCCTTTAGCACCCATAGCCACTGTTAAGCATGCAA 664
 DB 763 TGGCTGAAGAAATCAAAACCTTTTATGCAATTTAGCTCCAGTCGCACTGTGAAGTATACAA 704
 QY 665 AAGCCCCGGGACCAAAATTTTCTGCTGCAGATGATGATGATGATGATGATGATGATGATG 724
 DB 703 AAGCCCTTGAACAACTTAGGTTTATCTTCAACATGTTTCAAGATTTATTTGGTG 644
 QY 725 AAAAAAATTTCTGATCAGACCAAGATTTCTCAGCAACCTGTTTATTTTACCTTTGTGGCC 784
 DB 643 ACAAATATTCTACCCACACAATTTCTTTGATCAATTTCTTGGCACTCAAGTGTGTTCC 584
 QY 785 AGTGATTTCTGATCAGATTTTTCAGTGAATATGATGATGATGATGATGATGATGATGATG 844
 DB 583 GTGAGACACTGAATGTAATTTGCAAGCAATGCTTATTTATCATTTTGTGCTTTGACAGCG 524
 QY 845 ACAATATGAACATGAGCCGAGCAAGTGTATGCTGCCACACTCTTGTGGAACATCTG 904

Db	901	CTTCTGTTCAAAACATGTTCCATTGGACCCAGGCTGTTAAGTCTGGGAAATTCCAAGCTT	960
Qy	959	TTGACTGGGGAGTGAGACCAAAATCTGGAAAAATGCAATCAGCCAACTCCTGTAAGGT	1018
Db	961	ATGACTGGGGAAGCCCAAGTTGAGATAGGATGCACTATGATCAGTCCCAACCTCCCTACT	1020
Qy	1019	ACAGAGTCAGAGATATGACGGTCCCTACAGCAATGTGGACAGGAGGTGAGGACTGGGCTTT	1078
Db	1021	ACAATGTGACAGCCCATGAATGTACCAATTGCAGTGTGGACGTTGGCAAGGACCTGTTGG	1080
Qy	1079	CAATCCAGAGAGCGTGAATGCTGCTCTGTGAGGTGACCACTCATCTACCATTAAGA	1138
Db	1081	CTGACCCCAAGATGTTGGCCTTTTGTCTTCCAAAACCTCCCAATCTTATTTACCACAAGG	1140
Qy	1139	ATATTCTGGAATGGGCTCACGTGGATTTTCATCTGGGGTTTGGATGCTCCTCACCGTATGT	1198
Db	1141	AGATTCTCTTTTACAAATCACTGGACTTTATCTGGGCAATGGATGCCCTCAAGAAGTTT	1200
Qy	1199	ACAATGAATCATCCATCTGAT	1220
Db	1201	ACAATGACATTGTTTCTATGAT	1222

Search completed: February 19, 2004, 08:29:21
 Job time : 4672.38 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2004 Compugen Ltd.

OM nucleic - nucleic search, using sw model

Run on: February 19, 2004, 04:48:12 ; Search time 2775.48 Seconds
(without alignments)
11112.438 Million cell updates/sec

Title: US-10-042-431-46

Perfect score: 1269

Sequence: 1 atgttgaaacctgtcaag.....gacgggtgagccgattg 1269

Scoring table: IDENTITY NUC

Gapop 10.0 , Gapext 1.0

Searched: 22781392 seqs, 12152238056 residues

Total number of hits satisfying chosen parameters: 45562784

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database :

BST.*

1: em_estba.*

2: em_esthum.*

3: em_estin.*

4: em_estmu.*

5: em_estov.*

6: em_estpl.*

7: em_estro.*

8: em_htc.*

9: gb_estl.*

10: gb_est2.*

11: gb_htc.*

12: gb_est3.*

13: gb_est4.*

14: gb_est5.*

15: em_estfun.*

16: em_eston.*

17: em_gss_hum.*

18: em_gss_inv.*

19: em_gss_pln.*

20: em_gss_vrt.*

21: em_gss_fun.*

22: em_gss_mam.*

23: em_gss_mus.*

24: em_gss_pro.*

25: em_gss_rod.*

26: em_gss_phg.*

27: em_gss_vrl.*

28: gb_gss1.*

29: gb_gss2.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	961.8	75.8	2050	11 AK037091	AK037091 Mus muscu
2	961.8	75.8	2678	11 AK085719	AK085719 Mus muscu
3	960.2	75.7	2555	11 AK037214	AK037214 Mus muscu
4	618.2	48.7	2927	11 AK019504	AK019504 Mus muscu

5	505	39.8	954	10	BF135102	BF135102 601780142
6	497.6	39.2	1079	10	BF141567	BF141567 601788654
7	475.4	37.5	2138	11	AK088659	AK088659 Mus muscu
8	471.8	37.2	689	10	BF143546	BF143546 601790896
9	461.8	36.4	1350	11	AK010093	AK010093 Mus muscu
10	460.2	36.3	1347	11	AK010116	AK010116 Mus muscu
11	460.2	36.3	1349	11	AK009300	AK009300 Mus muscu
12	460.2	36.3	1350	11	AK010139	AK010139 Mus muscu
13	458.6	36.1	1340	11	AK010124	AK010124 Mus muscu
14	458.6	36.1	1343	11	AK009479	AK009479 Mus muscu
15	458.6	36.1	1345	11	AK009413	AK009413 Mus muscu
16	458.6	36.1	1346	11	AK009473	AK009473 Mus muscu
17	458.6	36.1	1346	11	AK009359	AK009359 Mus muscu
18	458.6	36.1	1346	11	AK009523	AK009523 Mus muscu
19	458.6	36.1	1346	11	AK009729	AK009729 Mus muscu
20	458.6	36.1	1348	11	AK010058	AK010058 Mus muscu
21	458.6	36.1	1349	11	AK009474	AK009474 Mus muscu
22	458.6	36.1	1349	11	AK009546	AK009546 Mus muscu
23	458.6	36.1	1349	11	AK009571	AK009571 Mus muscu
24	458.6	36.1	1349	11	AK009573	AK009573 Mus muscu
25	458.6	36.1	1349	11	AK010148	AK010148 Mus muscu
26	458.6	36.1	1350	11	AK009459	AK009459 Mus muscu
27	458.6	36.1	1350	11	AK009525	AK009525 Mus muscu
28	458.6	36.1	1350	11	AK009723	AK009723 Mus muscu
29	458.6	36.1	1350	11	AK010061	AK010061 Mus muscu
30	458.6	36.1	1350	11	AK010125	AK010125 Mus muscu
31	458.6	36.1	1350	11	AK010203	AK010203 Mus muscu
32	458.6	36.1	1351	11	AK009428	AK009428 Mus muscu
33	458.6	36.1	1352	11	AK009431	AK009431 Mus muscu
34	458.6	36.1	1356	11	AK010019	AK010019 Mus muscu
35	458.6	36.1	1356	11	AK010035	AK010035 Mus muscu
36	457.6	36.1	1346	11	AK009437	AK009437 Mus muscu
37	457.6	36.1	1347	11	AK010103	AK010103 Mus muscu
38	457.6	36.1	1349	11	AK010026	AK010026 Mus muscu
39	457	36.0	1346	11	AK009560	AK009560 Mus muscu
40	457	36.0	1347	11	AK010236	AK010236 Mus muscu
41	457	36.0	1348	11	AK010231	AK010231 Mus muscu
42	457	36.0	1349	11	AK009875	AK009875 Mus muscu
43	457	36.0	1349	11	AK009537	AK009537 Mus muscu
44	456	35.9	1349	11	AK010106	AK010106 Mus muscu
45	455.4	35.9	1345	11	AK009544	AK009544 Mus muscu

ALIGNMENTS

RESULT 1

AK037091

LOCUS

DEFINITION

AK037091 2050 bp mRNA linear

Mus musculus adult female vagina cDNA, RIKEN full-length enriched

library, clone:9930111004 product:BA30415.1 (NOVEL LIPASE)

(FRAGMENT) homolog [Homo sapiens], full insert sequence.

ACCESSION

AK037091.1 GI:26331937

VERSION

AK037091.1

KEYWORDS

HTC; CAP trapper.

SOURCE

Mus musculus (house mouse)

ORGANISM

Mus musculus

REFERENCE

1

Carninci, P., Shibata, Y., Hayatsu, N., Sugahara, Y., Shibata, K.,

Itoh, M., Kono, H., Okazaki, Y., Muramatsu, M. and Hayashizaki, Y.

Normalization and subtraction of cap-trapper-selected cDNAs to

prepare full-length cDNA libraries for rapid discovery of new genes

Genome Res. 10 (10), 1617-1630 (2000)

JOURNAL

MEDLINE

PUBMED

11042159

REFERENCE
AUTHORS

3 Shibata,K., Itoh,M., Aizawa,K., Nagaoka,S., Sasaki,N., Carninci,P.,
Konno,H., Akiyama,J., Nishi,K., Katsunai,T., Tashiro,H., Itoh,M.,
Sumi,N., Ishii,Y., Nakamura,S., Hazama,M., Nishine,T., Harada,A.,
Yamanoto,R., Matsumoto,H., Sakaguchi,S., Ikegami,T., Kashiwagi,K.,
Fujiwaka,S., Inoue,K., Togawa,Y., Izawa,M., Ohara,E., Watahiki,M.,
Yoneda,Y., Ishikawa,T., Ozawa,K., Tanaka,T., Matsuya,S., Kawai,J.,
Okazaki,Y., Muramatsu,M., Inoue,Y., Kira,A. and Hayashizaki,Y.
RIKEN integrated sequence analysis (RISA) system-384-format
sequencing pipeline with 384 multicapillary sequencer
Genome Res. 10 (11), 1757-1771 (2000)
20530913
11076861

REFERENCE
AUTHORS

4 Kawai,J., Shinagawa,A., Shibata,K., Yoshino,M., Itoh,M., Ishii,Y.,
Arakawa,T., Hara,A., Fukunishi,Y., Konno,H., Adachi,J., Fukuda,S.,
Aizawa,K., Izawa,M., Nishi,K., Kiyosawa,H., Kondo,S., Yamana,K.,
Saito,T., Okazaki,Y., Gojobori,T., Bono,H., Kasukawa,T., Saito,R.,
Kadota,K., Matsuda,H., Ashburner,M., Batalov,S., Casavant,T.,
Fleischmann,W., Gaasterland,T., Gissi,C., King,B., Kochiwa,H.,
Kuehl,P., Lewis,S., Matsuo,Y., Nikaide,I., Pesole,G.,
Quackenbush,J., Schriml,L.M., Staubli,F., Suzuki,R., Tomita,M.,
Wagner,L., Washio,T., Sawai,K., Okido,T., Furuno,M., Aono,H.,
Baldarelli,R., Barsh,G., Blake,J., Boiffelli,D., Bojunga,N.,
Carninci,P., de Bonaldo,M.F., Brownstein,M.J., Bult,C.,
Fletcher,C., Fujita,M., Gariboldi,M., Gustincich,S., Hill,D.,
Hofmann,M., Hume,D.A., Kamiya,M., Lee,N.H., Lyons,P.,
Marchionni,L., Mashima,J., Mazzarelli,J., Mombaerts,P., Nordone,P.,
Ring,B., Ringwald,M., Rodriguez,I., Sakamoto,N., Sasaki,H.,
Sato,K., Schonbach,C., Seya,T., Shibata,Y., Storch,K.F., Suzuki,H.,
Toyo-oka,K., Wang,K.H., Weitz,C., Whitaker,C., Wilming,L.,
Wyshaw-Boris,A., Yoshida,K., Hasegawa,Y., Kawaji,H., Kotsuki,S.
and Hayashizaki,Y.
Functional annotation of a full-length mouse cDNA collection
Nature 409 (6821), 685-690 (2001)
21085660
11217951

REFERENCE
AUTHORS

5 The FANTOM Consortium and the RIKEN Genome Exploration Research
Group Phase I & II Team.
Analysis of the mouse transcriptome based on functional annotation
of 60,770 full-length cDNAs
Nature 420, 563-573 (2002)

REFERENCE
AUTHORS

6 (bases 1 to 2050)
Adachi,J., Aizawa,K., Akimura,T., Arakawa,T., Bono,H., Carninci,P.,
Fukuda,S., Furuno,M., Hanagaki,T., Hara,A., Hashizume,W.,
Hayashida,K., Hayatsu,N., Hiramoto,K., Hiraoka,T., Hirozane,T.,
Hori,F., Imotani,K., Ishii,Y., Itoh,M., Kagawa,I., Kasukawa,T.,
Kato,H., Kawai,J., Kojima,Y., Kondo,S., Konno,H., Kouda,M.,
Koya,S., Kurihara,C., Matsuura,T., Miyazaki,A., Murata,M.,
Nakamura,M., Nishi,K., Nomura,K., Numazaki,R., Ohno,M., Ohsato,N.,
Okazaki,Y., Saito,R., Saitoh,H., Sakai,C., Sakai,K., Sakazume,N.,
Sano,H., Sasaki,D., Shibata,K., Shinagawa,A., Shiraki,T.,
Sogabe,Y., Tagami,T., Tagawa,A., Takahashi,P., Takaku-Akahira,S.,
Takeda,Y., Tanaka,T., Tomaru,A., Toya,T., Yasunishi,A.,
Muramatsu,M. and Hayashizaki,Y.
Direct Submission

REFERENCE
AUTHORS

Submitted (16-JUL-2001) Yoshihide Hayashizaki, The Institute of
Physical and Chemical Research (RIKEN), Laboratory for Genome
Exploration Research Group, RIKEN Genomic Sciences Center (GSC),
RIKEN Yokohama Institute, 1-7-22 Suehiro-cho, Tsurumi-Ku, Yokohama,
Kanagawa 230-0045, Japan (E-mail:genome-res@sc.riken.go.jp,
URL:http://genome.gsc.riken.go.jp/, Tel:81-45-503-9222,
Fax:81-45-503-9216)

COMMENT

cDNA library was prepared and sequenced in Mouse Genome
Encyclopedia Project of Genome Exploration Research Group in Riken
Genomic Sciences Center and Genome Science Laboratory in RIKEN.
Division of Experimental Animal Research in Riken contributed to
prepare mouse tissues.
Please visit our web site for further details.
URL:http://genome.gsc.riken.go.jp/
URL:http://fantom.gsc.riken.go.jp/
Location/Qualifiers

FEATURES

source	1. .2050	/organism="Mus musculus"
		/mol_type="mRNA"
		/strain="C57BL/6J"
		/db_xref="FANTOM DB:9930111004"
		/db_xref="taxon:10090"
		/clone="9930111004"
		/sex="female"
		/tissue_type="vagina"
		/clone_lib="RIKEN full-length enriched mouse cDNA library"
		/dev_stage="adult"
CDS	144..1412	/note="untranslated protein product; BA30415.1 (NOVEL LIPASE) (FRAGMENT) homolog (Homo sapiens) [SPTR Q96LG2, evidence: FASTY, 88.5%ID, 100%length, match=1119]
		/putative="putative"
		/codon_start=1
		/protein_id="BA29699.1"
		/db_xref="GI:26331938"
		/translation="MSBIIISRVVTVSHRVEIMLLIVLQNVNNSGHLPTKAADPE AFNVSEIIKHGYPSEVEYEDGYILSVNRIPRGQTLKKEGSRPVVLFQGLLG DSNWISNLPNSLIGFILLADAGDVGWNGSRGNTSRKHTLSIDDEFWAFSDMA RFDLPVINFILQKTSQKVVYVYSGQTMGFIAFTWPELAHKIMVIFALAPIATV KYAESPTKELLIPDMKIVLFGRELFCTPROLFYLQGMILDOICSNIIILL GGFNTNNNSRANVVAHPAGTSVQNLHWSQVNSBELRAFDWGSSETKQEKCNQ PTPIRYKVRDMVPTAMWTGGQWLSNPDVKTLLSEVNLIIHKNIPEWAHVDFIWG LDPQRYNEIILHMKQEPNLPQGTCTRVKL"
		/note="putative"
polya_signal	2027..2032	
polya_site	2050	
		/note="putative"
BASE COUNT	598 a 415 c 461 g 576 t	
ORIGIN		
	Query Match	75.8%; Score 961.8; DB 11; Length 2050;
	Best Local Similarity	85.4%; Pred. No. 1.5e-279;
	Matches 1084; Conservative	0; Mismatches 182; Indels 3; Gaps 1;
QY	1	ATGTTGGAAACCTTGTCAAGACAGTGGATTGTCTCACAGAAATGGAATGGCTTCG 60
DB	144	ATGTCAGAAATCTTGTCAAGAGTGTGGACTGTTTCGCACAGAGTGGAGATGGCTTCG 203
QY	61	ATTCTGGTGGCGTATATGTTCCAGAGAAATGTGAATTCAGTACATATGCCAACTAAAGCT 120
DB	204	ATTCTGTAGGTATTTACTTCCAGAGAAATGTGAATTCGGACATTTGCCACGAAAGCT 263
QY	121	GTGACCCAGAAAGCATTCATGAATATTAGTGAATCATCAATCAAGCTATTCCTGT 180
DB	264	GCGGATCCAGAAAGCATTCATGAATATTAGGGAATATCAATCAACACAAAGGGTTATCCAGT 323
QY	181	GAGGAATATGAAGTCGCAACTGAAGATGGGTATATCTTCTGTTAACAGATTCCTCGA 240
DB	324	GAGGATATGAAGTTGCAACGAAGATGGTATCATCTTCTGTGAACAGAAATCCCTCGG 383
QY	241	GGCCTAGTGCACCTTAAGAGACAGGTTCACAGCGCTGTGGTGTATCTGTCAGCATGGCCTA 300
DB	384	GGACAGACACCGTTTAAAGAGAGAGAGTCCAGGCGAGTGTGTACTGCAACATGGTCTT 443
QY	301	GTGAGAGTGTCTAGCAACTGATTTCCAACTGCCCAACATAGCTGGGCTTCATTCG 360
DB	444	CTTGGGATGTCTAGCAACTGATTTCCAACTGCCCAACATAGCTGGGCTTCATTCG 503
QY	361	GCAGATGTGGTGTGCTGATGGGAAACAGCAGGGGGAACAGCGCTGTCTCGAAAA 420
DB	504	GCAGATGTGGTGTGATGTGTGGATGGGAAACAGCAGGGGGAACAGCGCTGTCTCGGAAG 563
QY	421	CACAGACACTCTCCATAGACAAAGATGATTTGGGCTTTCAGTTATGATGAGATGGCT 480
DB	564	CACAGACCTCTCCATAGATCAAGATGATTTGGGCTTTCAGTTATGATGAGATGGCT 623
QY	481	AGGTTTGAACCTTCTCGCAGTGTATAAATTTATTTTTCAGAAAAACGGCCAGGAAAAAGATC 540
DB	624	AGGTTTGAACCTTCCAGCTGTGTATAAATTTATTTTTCAGAAAAACGGCCAGGAAAAAGATC 683

QY 541 TATTATCGGCTATTACAGGGGACCAACCATGGCTTTATTCGATTTTCCACCATGCCA 600
 Db |||||
 QY 684 TATTATCGGCTACTACAGGGGACCAACCATGGCTTTATTCGATTTTCCACCATGCCA 743
 Db |||||
 QY 601 GAGCTGGCTCAGAAATCAAAATGATTTTGGCTTTAGCACCCTAGCCTGTTAAGCAT 660
 Db |||||
 QY 744 GAGCTAGCTCATAAATCAAAATGATTTTGGCTTTAGCACCCTAGCCTGTTAATAT 803
 Db |||||
 QY 561 GCRAAAGCCCGGACCAATTTTGTCTGCGCAGATGATGATCAAGGATTTT 720
 Db |||||
 QY 804 GCAGAAGTCTGATCAAAATTTCTGCTGCGCAGATGATGATCAAGGATTTT 863
 Db |||||
 QY 721 GGCAAAAAGAAATTTCTGTATCAGACCAAGATTTCTCAGACAACTTGTATTACCTTTGT 780
 Db |||||
 QY 864 GGCAGACAGAGTTTATACCAAGACTAGATTTTTCAGACAGCTTTTATTACCTTTGT 923
 Db |||||
 QY 781 GGCAGAGTATCTGATCAGATTTGATGATATATCATGATTTCTGGTGTGATCAAC 840
 Db |||||
 QY 924 GGCAGATGATTTCTGACCAATCTGACCAACATCATCTTACTCTCTGGGAGGTTTAAAC 983
 Db |||||
 QY 841 ACCAAATATGACATGAGCGGACCAAGTATGATGCTGCCACACTCTTCTGCTGGAACA 900
 Db |||||
 QY 984 ACAACATATGACATGAGCGGACCAAGTATGATGCTGCCACACTCTTCTGCTGGAACG 1043
 Db |||||
 QY 901 TCTGTCAAAATATTTCTACTGAGCGGACCAAGTATGATGCTGCCACACTCTTCTGCTGGAAC 960
 Db |||||
 QY 1044 TCTGTGAGAAATCTCTCACTGAGCGGACCAAGTATGATGCTGCCACACTCTTCTGCTGGAAC 1103
 Db |||||
 QY 961 GACTGGGGAGTGAACCAAAATCTGAAATGCAATGCAATGCAATGCAATGCAATGCAATGCAAT 1020
 Db |||||
 QY 1104 GACTGGGGAGTGAACCAAAATCTGAAATGCAATGCAATGCAATGCAATGCAATGCAATGCAAT 1163
 Db |||||
 QY 1021 AGACTCAGAGATGACGGTCTCTTACAGCAATGTCAGAGAGTCTGAGTCTGAGTCTGAGTCTGAG 1080
 Db |||||
 QY 1164 AAGTTCGAGATATGATGTTCCCAACAGCAATGTCAGTCTGAGTCTGAGTCTGAGTCTGAGTCTGAG 1223
 Db |||||
 QY 1081 AATCCAGAGAGTGAATGCTCTCTGAGTGAACCACTATCTATCTATCTATCTATCTATCTATCTAT 1140
 Db |||||
 QY 1224 AATCCAGATGATGGAACCAATCTCTCTGAGTGAACCACTATCTATCTATCTATCTATCTATCTATCTAT 1283
 Db |||||
 QY 1141 ATTCTGATGAGTCTGAGTGAATTTCTCTGAGTGAATTTCTCTGAGTGAATTTCTCTGAGTGAATTTCTCT 1200
 Db |||||
 QY 1284 ATTCTGAGTGGTCTATGAGTGAATTTCTCTGAGTGAATTTCTCTGAGTGAATTTCTCTGAGTGAATTTCTCT 1343
 Db |||||
 QY 1201 AATGAATATCTATCTATCTATCTATCTATCTATCTATCTATCTATCTATCTATCTATCTATCTATCTAT 1260
 Db |||||
 QY 1344 AATGAATATCTATCTATCTATCTATCTATCTATCTATCTATCTATCTATCTATCTATCTATCTATCTAT 1400
 Db |||||
 QY 1261 GCCGTATTG 1269
 Db |||||
 QY 1401 GTCAATTTG 1409
 Db |||||
 RESULT 2
 LOCUS AK085719
 DEFINITION Mus musculus 10 days lactation, adult female mammary gland cDNA,
 RIKEN full-length enriched library, clone:D730026117
 product:BA30415.1 (NOVEL LIPASE) (FRAGMENT) homolog [Homo sapiens],
 full insert sequence.
 ACCESSION AK085719
 VERSION AK085719.1 GI:26351760
 KEYWORDS HTC; CAP trapper.
 SOURCE Mus musculus (house mouse)
 ORGANISM Mus musculus
 Eukaryota; Chordata; Craniata; Vertebrata; Euteleostomi;
 Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
 1
 REFERENCE Carninci, P. and Hayashizaki, Y.
 AUTHORS High-efficiency full-length cDNA cloning
 TITLE Mech. Enzymol. 303, 19-44 (1999)
 JOURNAL 99279253
 MEDLINE

10349636

PUBMED
REFERENCE
AUTHORS

Carninci, P., Shibata, Y., Hayatsu, N., Sugahara, Y., Shibata, K.,
 Itoh, M., Konno, H., Okazaki, Y., Muramatsu, M. and Hayashizaki, Y.
 Normalization and subtraction of cap-trapper-selected cDNAs to
 prepare full-length cDNA libraries for rapid discovery of new genes
 Genome Res. 10 (10), 1617-1630 (2000)

20499374

JOURNAL
MEDLINE
PUBMED

Shibata, K., Itoh, M., Aizawa, K., Nagaoka, S., Sasaki, N., Carninci, P.,
 Konno, H., Akiyama, J., Nishi, K., Kitsuana, T., Tashiro, H., Itoh, M.,
 Sumi, N., Ishii, Y., Nakamura, S., Hazama, M., Nishine, T., Harada, A.,
 Yamamoto, R., Matsumoto, H., Sakaguchi, S., Ikegami, T., Kashiwagi, K.,
 Fujiwara, S., Inoue, K., Togawa, Y., Izawa, M., Ohara, E., Watabiki, M.,
 Yoneda, Y., Ishikawa, T., Ozawa, K., Tanaka, T., Matsura, S., Kawai, J.,
 Okazaki, Y., Muramatsu, M., Inoue, Y., Kira, A. and Hayashizaki, Y.
 RIKEN integrated sequence analysis (RISA) system--384-format
 sequencing pipeline with 384 multicapillary sequencer
 Genome Res. 10 (11), 1757-1771 (2000)

20530913

JOURNAL
MEDLINE
PUBMED

Kawai, J., Shinagawa, A., Shibata, K., Yoshino, M., Itoh, M., Ishii, Y.,
 Arakawa, T., Hara, A., Fukunishi, Y., Konno, H., Adachi, J., Fukuda, S.,
 Aizawa, K., Izawa, M., Nishi, K., Kiyosawa, H., Kondo, S., Yamana, I.,
 Saito, T., Okazaki, Y., Gojobori, T., Bono, H., Kasukawa, T., Saito, R.,
 Kadota, K., Matsuda, H., Ashburner, M., Batalov, S., Casavant, T.,
 Fleisemann, W., Gaasterland, T., Gissi, C., King, B., Kochiwa, H.,
 Kuehl, P., Lewis, S., Matsuo, Y., Nikaide, I., Pesole, G., Tomita, M.,
 Quackenbush, J., Schriml, L.M., Staib, F., Suzuki, R., Anon, H.,
 Wagner, L., Washio, T., Sakai, K., Okido, T., Furuno, M., Aono, M.,
 Baldarelli, R., Barsh, G., Blake, J., Boffelli, D., Bojunga, N.,
 Carninci, P., de Bernaldo, M.P., Brownstein, M.J., Bult, C.,
 Fletcher, C., Fujita, M., Gariboldi, M., Gustincich, S., Hill, D.,
 Hofmann, M., Hume, D.A., Kamiya, M., Lee, N.H., Lyons, P.,
 Marchionni, L., Mashima, J., Mazzarelli, J., Mombauts, P., Nordone, P.,
 Ring, B., Ringwald, M., Rodriguez, I., Sakamoto, N., Sasaki, H.,
 Sato, K., Schonbach, C., Seya, T., Shibata, Y., Storch, K.F., Suzuki, H.,
 Toyooka, K., Wang, K.H., Weitz, C., Whittaker, C., Wilming, L.,
 Wynshaw-Boris, A., Yoshida, K., Hasegawa, Y., Kawaji, H., Kohtsuki, S.
 and Hayashizaki, Y.
 Functional annotation of a full-length mouse cDNA collection
 Nature 409 (6821), 685-690 (2001)

21085660

JOURNAL
MEDLINE
PUBMED

The FANTOM Consortium and the RIKEN Genome Exploration Research
 Group Phase I & II Team.
 Analysis of the mouse transcriptome based on functional annotation
 of 60,770 full-length cDNAs
 Nature 420, 563-573 (2002)

6 (bases 1 to 2678)

JOURNAL
REFERENCE
AUTHORS

Adachi, J., Aizawa, K., Akimura, T., Arakawa, T., Bono, H., Carninci, P.,
 Fukuda, S., Furuno, M., Hanagaki, T., Hara, A., Hashizume, M.,
 Hayashida, K., Hayatsu, N., Hiramoto, K., Hiraoka, T., Hirozane, T.,
 Hori, F., Imotani, K., Ishii, Y., Itoh, M., Kagawa, I., Kasukawa, T.,
 Kato, H., Kawai, J., Kojima, Y., Kondo, S., Konno, H., Kouda, M.,
 Koya, S., Kurahara, C., Matsuyama, T., Miyazaki, A., Murata, M.,
 Nakamura, M., Nishi, K., Nomura, K., Numazaki, R., Ohnishi, N.,
 Okazaki, Y., Saito, R., Saitoh, H., Sakai, C., Sakai, K., Sakazume, N.,
 Sano, H., Sasaki, D., Shibata, K., Shinagawa, A., Shiraki, T.,
 Sogabe, Y., Tagami, M., Tagawa, A., Takahashi, F., Takaku-Akahira, S.,
 Takada, Y., Tanaka, T., Tomaru, A., Toyota, T., Yasunishi, A.,
 Muramatsu, M. and Hayashizaki, Y.

Direct Submission

JOURNAL
TITLE

Submitted (16-APR-2002) Yoshihide Hayashizaki, The Institute of
 Physical and Chemical Research (RIKEN), Laboratory for Genome
 Exploration Research Group, RIKEN Genomic Sciences Center (GSC),
 RIKEN Yokohama Institute, 1-7-22 Suehiro-cho, Tsurumi-ku, Yokohama,
 Kanagawa 230-0045, Japan [E-mail: genome-res@gscc.riken.go.jp,
 URL: http://genome.gsc.riken.go.jp/, Tel: 81-45-503-9222,
 Fax: 81-45-503-9216]

COMMENT

cDNA library was prepared and sequenced in Mouse Genome Encyclopedia Project of Genome Exploration Research Group in Riken Genomic Sciences Center and Genome Science Laboratory in RIKEN. Division of Experimental Animal Research in Riken contributed to prepare mouse tissues.

Please visit our web site for further details.

URL: <http://genome.gsc.riken.go.jp/>

URL: <http://fantom.gsc.riken.go.jp/>

FEATURES

source

```

1..2678
/organism="Mus musculus"
/mol_type="mRNA"
/strain="C57BL/6J"
/db_xref="PANTOM DB:D730026117"
/db_xref="taxon:10090"
/clone="D730026117"
/sex="female"
/tissue type="mammary gland"
/clone lib="RIKEN full-length enriched mouse cDNA library"
/dev_stage="10 days lactation, adult"
118..1386
/note="unmated protein product; BA30415.1 (NOVEL LIPASE) (FRAGMENT) homolog [Homo sapiens] (SPR|O96LG2, evidence: FASTV, 88.5%id, 100%length, match=1119) putative"
/codon_start=1
/protein_id="BAC39517.1"
/db_xref="GI:26351761"
/translation="MSEILSRVWTVSHRVEIWLILVAILQNVNSGHLPTKAADPE AFNVSEIHKHGYDSEEVATEDGYILSVNRI PRGQRLKKGSRVVLQHLGLG DASNI SNLPNLSGFI LADAGFDVWNGSRNTWSRKHKLSDQDSEFWAFSDENMA RFDLPVINFILQTKQKVVYVGSQGTTFMGFIASFMPDLAKHKNYFALAPIATV KVARSPGTFLLPDMIKVLPGRQEFYQTFRQLFIYLCGMILDOI GSNILILL KGFNTNMMSRANVVAHTPAGTSVQNLHWSOAVNSGELRAPDWGSETKNQKCNQ PPPIKVPZDMVPTAMWVGQDLSNPDDVKTLLSEVNTLYHKNIPFWAHVDFING LPAQKVYNEIHLMKQENLPFGICRKL"
2657..2662
/note="putative"
2678
/note="putative"

```

CDS

```

1..2678
/codon_start=1
/protein_id="BAC39517.1"
/db_xref="GI:26351761"
/translation="MSEILSRVWTVSHRVEIWLILVAILQNVNSGHLPTKAADPE AFNVSEIHKHGYDSEEVATEDGYILSVNRI PRGQRLKKGSRVVLQHLGLG DASNI SNLPNLSGFI LADAGFDVWNGSRNTWSRKHKLSDQDSEFWAFSDENMA RFDLPVINFILQTKQKVVYVGSQGTTFMGFIASFMPDLAKHKNYFALAPIATV KVARSPGTFLLPDMIKVLPGRQEFYQTFRQLFIYLCGMILDOI GSNILILL KGFNTNMMSRANVVAHTPAGTSVQNLHWSOAVNSGELRAPDWGSETKNQKCNQ PPPIKVPZDMVPTAMWVGQDLSNPDDVKTLLSEVNTLYHKNIPFWAHVDFING LPAQKVYNEIHLMKQENLPFGICRKL"
2657..2662
/note="putative"
2678
/note="putative"

```

polyA_site

```

2678
/note="putative"

```

polyA_site

```

2678
/note="putative"

```

BASE COUNT 795 a 522 c 555 g 806 t

ORIGIN

```

Query Match 75.8%; Score 961.8; DB 11; Length 2678;
Best Local Similarity 85.4%; Pred. No. 1.8e-279;
Matches 1084; Conservative 0; Mismatches 182; Indels 3; Gaps 1;

Qy 1 ATGTGGAACCTTGTCACAGACAGTGGATTGTCTCACAGAGATGGAATGGGCTTCG 60
Db 118 ATGTGGAATCTTGTCACAGAGTGGACTGTGTCACAGAGTGGAGATGGCTTCG 177

Qy 61 ATTCTGGTGGCTATATGTTCCAGAGAAATGTGAATTCAGTACATATGCCAACTAAGCT 120
Db 178 ATTCTGGTGGCTATATGTTCCAGAGAAATGTGAATTCAGTACATATGCCAACTAAGCT 237

Qy 121 GTGACCCAGACAGCTTACATATATGTTGAATTCATCCATCCATCAAGCTATCCCTGT 180
Db 238 GCGGATCCAGAGATTCATGAAATGTAGGAAATCATCAACCAAGGTTATCCAGT 297

Qy 181 GAGGAATATGAAGTCGCAATCTGAAGATGGGTATATCTTCTGTGTTAAGGATTCCTCGA 240
Db 298 GAGGATATGAAGTTGCAACGGAAGATGGGTATATCTTCTGTGTTAAGGATTCCTCGG 357

Qy 241 GGCTAGTGCAACTAAGAGACAGGTTCCAGGCTGTGGTGTACTGCAGCATGSCCTA 300
Db 358 GGACAGACACGGTTAAAGAGAGGAGGATCCAGGCCAGTGGTGTACTGCAACATGGCTTT 417

Qy 301 GTTGGAGGTGTAGCACTGGATTCCAACTGCTCCCAACAAATAGCTGGGCTTCATCTG 360
Db 418 CTTGGGATGCTAGCACTGGATTCCAACTGCTCCCAACAAATAGCTGGGCTTCATCTG 477

Qy 361 GCAGATCTGTTTTGACGTGTGGATGGGGAACAGCAGGGGAAACGCTGCTCGAAA 420

```

```

Db 478 GCAGATGCAGGTTTTCATGTTGGATGGGAAACACAGAGGAAACACCTGGTCTCGGAAG 537
Qy 421 CACAAGACACTCTCCATAGACCAAGATGAGTCTTGGGCTTCAGTATGATGATGAGTGGCT 480
Db 538 CACAAGACCTCTCCATAGATCAAGATGAGTCTTGGGCTTCAGTATGATGATGAGTGGCT 597
Qy 481 AGGTTTGAACCTCTCTGAGTGAATAAATTTATTTTTCAGAAAACCGGCGCAGGAAAGATC 540
Db 598 AGGTTTGAACCTCTCTGAGTGAATAAATTTATTTTTCAGAAAACCGGCGCAGGAAAGATC 657
Qy 541 TATTATGTCGGCTATTCACAGGGCACCAATGGGCTTTTATGCAATTTTCCACCAATGCCA 600
Db 658 TATTATGTCGGCTATTCACAGGGCACCAATGGGCTTTTATGCAATTTTCCACCAATGCCA 717
Qy 601 GAGCTGCTCAGAAAATCAAAATGATTTTGTAGCACCACCATAGCCACTGTTAAGCAT 660
Db 718 GAGCTAGCTCATTAATCAAAATGATTTTGTAGCTTCTATAGCCACTGTTAATAT 777
Qy 661 GCAAAAGCCCGGGACCAAAATTTTGTGTCGCCAGATATGATGATCAAGGATGTTT 720
Db 778 GCAAAAGCTCCTGGTACCAAAATTTCTGCTGCTGCCAGATATGATGATCAAGGATGTTT 837
Qy 721 GGCAAAAGAAATTTCTGATCAGACCAATTTCTCAGACCAACTTGTATTTTACCTTCT 780
Db 838 GGCAGACAGAGTTTTTATACCAAGATGATTTTTCAGACAGTTTTTATTTACCTTCT 897
Qy 781 GGCAGAGTATTTCTGATCAGATTTTCTAGTAATATCATTTTCTGCGGTGGATTCAC 840
Db 898 GGCAGAGATTTCTGACCAAAATCTGCAGCAACATCATCTTCTCTGGGAGGTTTTAAC 957
Qy 841 ACCAAATATGAAATGAGCCGAGCAAGTATATGCTGCCACACTCTTCTGCGAACA 900
Db 958 ACAAAATATGAAATGAGCCGAGCAAGTATGCTGCCCATATACACTGCTGGAACG 1017
Qy 901 TCTGTGCAAAATATTTCTACACTGAGCCAGGCAAGTGAATTTCTGGTGAATCTCCGGGCA 960
Db 1018 TCTGTGCAAGCAATTTCTCACTGAGCCAGGCGGTGAATTTCTGGGGAATCTTGTGCTTT 1077
Qy 961 GACTGGGGAGTGAGACCAAAATCTGAAAATATGCAATCAGCAACTCTCTGTAAGTAC 1020
Db 1078 GACTGGGGAAGTGAGACCAAAATCTGAGGAAATGTAAATCAGCAACTCTCTATAAGTAC 1137
Qy 1021 AGAGTCAGAGATATGACGGTCCCTCAGCAATGTGACAGGAGGTGAGGACTGGCTTTCA 1080
Db 1138 AAGTTTCAGATATGATGTTCCCAACAGCAATGTGAGTGGAGTCAAGACTGGCTTTCA 1197
Qy 1081 AATCCAGACAGCTGGAATGCTGCTCTCTGAGGTGACCAACCTCATCTACCAAGAAC 1257
Db 1198 AATCCAGATGATGGAACCAATTAATTTCTGAAAGTAAACCAACCTCATCTACCAAGAAC 1257
Qy 1141 ATTCTGAAATGGGCTCAGTGGATTTTCATCTGGGTTTGGATGCTCTCTACCGTATGTAC 1200
Db 1258 ATTCTGAGTGGGCTCATGTTGATTTCACTGGGGCTGGATGCCCTCAGCGTGTTCAC 1317
Qy 1201 AATGAAATCATCCATCTGATGTCAGAGGAGAGACCAACTTTTCCAGGGAGCGGTGTAG 1260
Db 1318 AATGAAATTTATCATCTGATGATGAGC---AGGAGCCCAACCTTCCCGGGAACCTGCAGG 1374
Qy 1261 GCCGTATTG 1269
Db 1375 GTCAAAATTG 1383

```

RESULT 3

AK037214

LOCUS

DEFINITION

AK037214

ACCESSION

VERSION

KEYWORDS

SOURCE

AK037214 2555 bp mRNA linear HTC 05-DEC-2002
Mus musculus 6 days neonate skin cDNA, RIKEN full-length enriched
library, clone:AK030010N08 product:BA30415.1 (NOVEL LIPASE)
(FRAGMENT) homolog [Homo sapiens], full insert sequence.

AK037214.1 GI:26332053

HTC; CAP trapper.

Mus musculus (house mouse)

ORGANISM	Mus musculus	TITLE	Direct Submission
REFERENCE	Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.	JOURNAL	Submitted (16-JUL-2001) Yoshihide Havaehizaki, The Institute of Physical and Chemical Research (RIKEN), Laboratory for Genome Exploration Research Group, RIKEN Genomic Sciences Center (GSC), RIKEN Yokohama Institute, 1-7-22 Suehiro-cho, Tsurumi-ku, Yokohama, Kanagawa 230-0045, Japan (E-mail: genome-res@gsc.riken.go.jp, URL: http://genome.gsc.riken.go.jp/, Tel: 81-45-503-9222, Fax: 81-45-503-9216)
AUTHORS	Carninci, P. and Hayashizaki, Y.	COMMENT	cDNA library was prepared and sequenced in Mouse Genome Encyclopedia Project of Genome Exploration Research Group in Riken Genomic Sciences Center and Genome Science Laboratory in RIKEN. Division of Experimental Animal Research in Riken contributed to prepare mouse tissues.
JOURNAL	High-efficiency full-length cDNA cloning	URL: http://genome.gsc.riken.go.jp/	
MEDLINE	Meth. Enzymol. 303, 19-44 (1999)	URL: http://fantom.gsc.riken.go.jp/	
PUBMED	99279253	Location/Qualifiers	
REFERENCE	10349636	1. 2555	
AUTHORS	Carninci, P., Shibata, Y., Hayatsu, N., Sugahara, Y., Shibata, K., Itoh, M., Konno, H., Okazaki, Y., Muramatsu, M. and Hayashizaki, Y.	/organism="Mus musculus"	
TITLE	Normalization and subtraction of cap-trapper-selected cDNAs to prepare full-length cDNA libraries for rapid discovery of new genes	/mol_type="mRNA"	
JOURNAL	Genome Res. 10 (10), 1617-1630 (2000)	/strains="C57BL/6J"	
MEDLINE	20499374	/db_xref="FANTOM_DB:A030010N08"	
PUBMED	11042159	/db_xref="taxon:10090"	
REFERENCE	3	/clone="A030010N08"	
AUTHORS	Shibata, K., Itoh, M., Aizawa, K., Nagao, K., Sasaki, N., Carninci, P., Konno, H., Akiyama, J., Nishi, K., Kitsuai, T., Tashiro, H., Itoh, M., Sumi, N., Ishii, Y., Nakamura, S., Hazama, M., Nishine, T., Harada, A., Yamamoto, R., Matsumoto, H., Sakaguchi, S., Ikegami, T., Kashiwagi, K., Fujiwaka, S., Inoue, K., Togawa, Y., Izawa, M., Ohara, E., Watabiki, M., Yoneda, Y., Ishikawa, T., Ozawa, K., Tanaka, T., Matsuura, S., Kawai, J., Okazaki, Y., Muramatsu, M., Inoue, Y., Kira, A. and Hayashizaki, Y.	/tissue_type="skin"	
TITLE	RIKEN integrated sequence analysis (RISA) system--384-format sequencing pipeline with 384 multiplexed sequencer	/clone_lib="RIKEN full-length enriched mouse cDNA library"	
JOURNAL	Genome Res. 10 (11), 1757-1771 (2000)	/dev_stage="6 days neonate"	
MEDLINE	20530913	126. 1394	
PUBMED	11076861	/note="unnamed protein product; BA30415.1 (NOVEL LIPASE) (FRAGMENT) homolog [Homo sapiens] (SPTR Q96LG2, evidence: FASTA, 88.5%ID, 100%length, match=1119)	
REFERENCE	4	putative"	
AUTHORS	Kawai, J., Shinagawa, A., Shibata, K., Yoshino, M., Itoh, M., Ishii, Y., Arakawa, T., Hara, A., Fukunishi, Y., Konno, H., Adachi, J., Fukuda, S., Aizawa, K., Izawa, M., Nishi, K., Kiyosawa, H., Kondo, S., Yamana, I., Saito, T., Okazaki, Y., Gojobori, T., Bono, H., Kasukawa, T., Saito, R., Kadota, K., Matsuda, H., Asbuerger, M., Batalov, S., Casavant, T., Fleischmann, W., Gaasterland, T., Gissi, C., King, B., Kochiwa, H., Kuehl, P., Lewis, S., Matsuo, Y., Nikaido, I., Pesole, G., Kachiwa, H., Quackenbush, J., Schrim, L.M., Staubli, F., Suzuki, G., Tomita, M., Wagner, L., Washio, T., Sakai, K., Okido, T., Furuno, M., Aono, H., Baldaire, R., Barsh, G., Blake, J., Boffelli, D., Bojunga, N., Carninci, P., de Bonaldo, M.F., Brownstein, M.J., Bult, C., Fletcher, C., Fujita, M., Garibol, M., Gustincich, S., Hill, D., Hofmann, M., Hume, D.A., Kaniwa, M., Lee, N.H., Lyons, P., Marchionni, L., Mashima, J., Mazzarelli, J., Mombaerts, P., Nordone, P., Ring, B., Ringwald, M., Rodriguez, I., Sakamoto, N., Sasaki, H., Sato, K., Schonbach, C., Seya, T., Shibata, Y., Storch, K.F., Suzuki, H., Toyo-oka, K., Wang, K.H., Weitz, C., Wittaker, C., Wilming, L., Wynshaw-Boris, A., Yoshida, K., Hasegawa, Y., Kawaji, H., Kohtsuki, S. and Hayashizaki, Y.	/codon_start=1	
TITLE	Functional annotation of a full-length mouse cDNA collection	/protein_id="BAC29757.1"	
JOURNAL	Nature 409 (6821), 685-690 (2001)	/db_xref="GI:26332054"	
MEDLINE	21085660	/translation="MSEILSRVTVSHRVEIWLILVAYLLQNVNSHLPTKAADPE AFNVSEIHKHGVSEVEVATEDGVLSVNRIPRGOTRLKKGSGSPVVLQHLGLG DASWISNLPNSLGFILADAFVDMNSRGNTWSRKHLSIDQDEWAFVSDEMA RFDPAVNFILQTKQKYYIGVYSGTQINGTAFSPMLAHKIKNYFALAIATV KYASPGTKFLLPDMIMKVLFGQEFYQFRPQLFYLCGGMILQDCISNILL GGFNNNNRANRYVAHTPAGTSLHWSQVNSGELRAFDWGSSTKNOBCKNQ PTTRYKVRDMVPTAMVTGGQDWSLNPDDVTKLLSEVTLNLYHKNIPEWAHVDPIWG LDAPORVYNEIHLMKOBPNLPQGTQCRVKL"	
PUBMED	11217851	BASE COUNT	752 a 501 c 536 g 766 t
REFERENCE	5	ORIGIN	
AUTHORS	The FANTOM Consortium and the RIKEN Genome Exploration Research Group Phase I & II Team.	Query Match	75.7%; Score 960.2; DB 11; Length 2555;
TITLE	Analysis of the mouse transcriptome based on functional annotation of 60,770 full-length cDNAs	Best Local Similarity	85.3%; Pred. No. 5.4e-279;
JOURNAL	Nature 420, 563-573 (2002)	Matches 1083; Conservative	0; Mismatches 183; Indels 3; Gaps 1;
MEDLINE	11217851		
PUBMED	1217851		
REFERENCE	6		
AUTHORS	Adachi, J., Aizawa, K., Akimura, T., Arakawa, T., Bono, H., Carninci, P., Fukuda, S., Furuno, M., Hanagaki, T., Hara, A., Hashizume, M., Hayashida, K., Hayatsu, N., Hiramoto, K., Hiraoka, T., Hirozane, T., Hori, P., Imotani, K., Ishii, Y., Itoh, M., Kagawa, I., Kasukawa, T., Katoh, H., Kawai, J., Kojima, Y., Kondo, S., Konno, H., Kouda, M., Koya, S., Kurthara, C., Matsuyama, T., Miyazaki, A., Murata, M., Nakamura, M., Nishi, K., Nomura, K., Numazaki, R., Ohno, M., Ohsato, N., Okazaki, Y., Saito, R., Saich, H., Sakai, C., Sakai, K., Sakazume, N., Sano, H., Sasaki, D., Shibata, K., Shinagawa, A., Shiraki, T., Sogabe, Y., Tagami, M., Tagawa, A., Takahashi, F., Takaku-Akahira, S., Takeda, Y., Tanaka, T., Tomaru, A., Toya, T., Yasunishi, A., Muramatsu, M. and Hayashizaki, Y.	QY	1 ATGTTGGAAACCTTCTCAAGACAGTGGATTCTCTCACACAGATGGAATCTGCTTCTG 60
TITLE	Analysis of the mouse transcriptome based on functional annotation of 60,770 full-length cDNAs	Db	126 ATGTGAGAAATCTTGTCAAGAGTGGGACTGTTTCCACAGAGTGGAGATATGCTTCTG 185
JOURNAL	Nature 420, 563-573 (2002)	QY	61 ATTCTGGTGGCGTATATGTTCCAGAGAAATGGAATTCAGTACATATATCCCAACTAAAGCT 120
MEDLINE	11217851	Db	186 ATTCTGGTAGCGTATTTACTCCAAAGAAATGTGAATCGGACATTTTCCCAACCAAGCT 245
PUBMED	1217851	QY	121 GTGGACCCAGAGCAATTCATGATATAGTGAATTCATCAATCCCAATCAAGGCTATCCCTGT 180
REFERENCE	7	Db	246 GCGGATCCAGAGCAATTCATGATATAGTGAATTCATCAATCCCAATCAAGGCTATCCCAAGT 305
AUTHORS	The FANTOM Consortium and the RIKEN Genome Exploration Research Group Phase I & II Team.	QY	181 GAGGATATGAGTCCGCACTGAAGATGGGTATATCTCTTCTGTTAAACAGGATTCCTCGA 240
TITLE	Analysis of the mouse transcriptome based on functional annotation of 60,770 full-length cDNAs	Db	306 GAGGATATGAGTTCGCAAGAGTGGTATCTCTTCTGTTAAACAGGATTCCTCGG 365
JOURNAL	Nature 420, 563-573 (2002)	QY	241 GGCCTAGTGCAACCTAAGAGCAGGTTCCAGGCTCTGGTGTGTACTCTGACAGATGGCCTA 300
MEDLINE	11217851	Db	366 GGACACACACCGTCTAAAGAGAGAGGATCCAGGCGAGTGGTGTGTACTCTGCAACATGGTCT 425
PUBMED	1217851	QY	301 GTTGGAGGTGTAGCAACTGGATTTCCAACTGCCCAACAATAGCTGGGCTTCTATTCTG 360
REFERENCE	8	Db	426 CTTGGGATGTAGCAACTGGATTTCCAACTGCCCAACAATAGCTGGGCTTCTATTCTG 485

QY 361 GCAGATGCTGTTTTCAGCTGCGATGGGACACAGCGGGAACGCTGCTCGAAA 420
 Db 486 GCAGATGCGAGTTTTCAGTGTGGATGGAAACAGCAGAGGAAACACCTGCTCGGAAG 545
 QY 421 CACAAGACACTCCATAGACCAAGATGAGTTCTGGGCTTTTCAGTTATGATGAGATGGCT 480
 Db 546 CACAAGACCTCTCCATAGATCAAGATGAGTTCTGGGCTTTTCAGTTATGATGAGATGGCT 605
 QY 481 AGTTTTCAGCTTCTGCGATGATAAATCTTTATTTTGCAGAAACCGGGCCAGGAAAGATC 540
 Db 606 AGTTTTCAGCTTCTGCGATGATAAATCTTTATTTTGCAGAAACCGGGCCAGGAAAGATC 665
 QY 541 TATTATGTCGGCTTTCAGCGGACCAACCATGGCTTTTATGCAATTTTCCACCATGCGCA 600
 Db 666 TATTATGTCGGCTTTCAGCGGACCAACCATGGCTTTTATGCAATTTTCCACCATGCGCA 725
 QY 601 GAGTGGCTCAGAAATCAAAATGATTTTGTGTTTGTAGCACCATGACCACTGTTAAGCAT 660
 Db 726 GAGTGGCTCAGAAATCAAAATGATTTTGTGTTTGTAGCACCATGACCACTGTTAAGCAT 785
 QY 661 GCAAAAGCCCGGACCAAAATTTTGTGTTTGTGTTTGTGTTTGTGTTTGTGTTTGTGTTT 720
 Db 786 GCAAGAGTCTGTTTACCAAAATTTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 845
 QY 721 GCAAAAGAGATTTCTGTATCAGACCAATTTCTCAGCAAACTTTGTTTATTTACCTTTGT 780
 Db 846 GGCAGACAGAGTTTATATACCAAGATGATTTTTCAGACAGCTTTTATTTTACCTTTGT 905
 QY 781 GGCAGGATGATTTTATCAGATTTTGTATGATATCATGTTTCTGTTTCTGTTTCTGTTTCT 840
 Db 906 GGCAGATGATTTTATGACCAATCTGACCAATCTGACCAATCTGACCAATCTGACCAATCT 965
 QY 841 ACCAACAATATGAACATGACCGGACCAAGTGTATATGTTGCTGCCACACTCTTCTGCGAACA 900
 Db 966 ACAACAATATGAACATGACCGGACCAAGTGTATATGTTGCTGCCACACTCTTCTGCGAACA 1025
 QY 901 TCTGTGCAAAATTTCTACCTGACCGGACCAAGTGTATATGTTGCTGCCACACTCTTCTGCGAACA 960
 Db 1026 TCTGTGCAAAATTTCTACCTGACCGGACCAAGTGTATATGTTGCTGCCACACTCTTCTGCGAACA 1085
 QY 961 GACTGGGGAGTGCAGACCAAAATCTGGAATAATGCAATCAGCAACTCTTCTGCGAACA 1020
 Db 1086 GACTGGGGAGTGCAGACCAAAATCTGGAATAATGCAATCAGCAACTCTTCTGCGAACA 1145
 QY 1021 AGAGTCAGATATGACGCTCCCTAAGCAATGTCAGAGAGGTCAGAGTGGCTTCTCA 1080
 Db 1146 AAGATTCAGATATGATGTTGTCACCAAGCAATGTCAGAGTGGCTTCTCA 1205
 QY 1081 AATCCAGAGACGTGAAATGCTCTCTGAGGTGACCAACTCATCTACCATTAAGAT 1140
 Db 1206 AATCCAGATGATGTAACCAATTTCTGAGTAACTCACTCACTCACTCACTCACTCACTCA 1265
 QY 1141 ATTCTGTAATGCTCAGTGGATTTCTCTGAGGTGACCAACTCATCTACCATTAAGAT 1200
 Db 1266 ATTCTGTAATGCTCAGTGGATTTCTCTGAGGTGACCAACTCATCTACCATTAAGAT 1325
 QY 1201 AATGAATATCATCTGATGACGAGGAGGACCAACTCTTCCAGGACGGTGTGAG 1260
 Db 1326 AATGAATATCATCTGATGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 1382
 QY 1261 GCCGTATTG 1269
 Db 1383 GTCAAAATTG 1391

RESULT 4
 AK019504
 LOCUS
 DEFINITION Mus musculus 0 day neonate skin cDNA, RIKEN full-length enriched library, clone:4632427C23 product:BA30415.1 (NOVEL LIPASE) (FRAGMENT) homolog [Homo sapiens], full insert sequence.
 ACCESSION AK019504

VERSION
 KEYWORDS
 SOURCE
 ORGANISM
 REFERENCE
 AUTHORS
 TITLE
 JOURNAL
 MEDLINE
 PUBMED
 REFERENCE
 AUTHORS
 TITLE
 JOURNAL
 MEDLINE
 PUBMED
 REFERENCE
 AUTHORS
 TITLE
 JOURNAL
 MEDLINE
 PUBMED
 REFERENCE
 AUTHORS

AK019504.1 GI:12859754
 HTC; CAP trapper
 Mus musculus (house mouse)
 Mus musculus
 Mammalia; Euthera; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
 1 Carninci, P. and Hayashizaki, Y.
 High-efficiency full-length cDNA cloning
 Meth. Enzymol. 303, 19-44 (1999)
 99279253
 10349636
 2 Carninci, P., Shibata, Y., Hayatsu, N., Sugahara, Y., Shibata, K., Itoh, K., Konno, H., Okazaki, Y., Muramatsu, M. and Hayashizaki, Y.
 Normalization and subtraction of cap-trapper-selected cDNAs to prepare full-length cDNA libraries for rapid discovery of new genes
 Genome Res. 10 (10), 1617-1630 (2000)
 20499374
 11042159
 3 Shibata, K., Itoh, M., Aizawa, K., Nagaoka, S., Sasaki, N., Carninci, P., Konno, H., Akiyama, J., Nishi, K., Kishimoto, T., Tashiro, H., Itoh, M., Sumi, N., Ishii, Y., Nakamura, S., Hazama, M., Nishine, T., Harada, A., Yamamoto, R., Matsumoto, H., Sakaguchi, S., Ikegami, T., Kashiwagi, K., Fujiwara, S., Inoue, K., Togawa, Y., Izawa, M., Ohara, E., Watahiki, M., Yoneda, Y., Ishikawa, T., Ozawa, K., Tanaka, T., Matsura, S., Kawai, J., Okazaki, Y., Muramatsu, M., Inoue, Y., Kira, A. and Hayashizaki, Y.
 RIKEN integrated sequence analysis (RISA) system--384-format sequencing pipeline with 384 multicapillary sequencer
 Genome Res. 10 (11), 1757-1771 (2000)
 20530913
 11076861
 4 Kawai, J., Shinagawa, A., Shibata, K., Yoshino, M., Itoh, M., Ishii, Y., Akakawa, T., Hara, A., Fukunishi, Y., Konno, H., Adachi, J., Fukuda, S., Aizawa, K., Izawa, M., Nishi, K., Kiyosawa, H., Kondo, S., Yamanaka, I., Saito, T., Okazaki, Y., Gojobori, T., Bono, H., Kasukawa, T., Saito, R., Kadota, K., Matsuda, H., Ashburner, M., Batalov, S., Casavant, T., Fleischmann, W., Gaasterland, T., Gissi, C., King, B., Kochiwa, H., Kuehl, P., Lewis, S., Matsuo, Y., Nikaide, I., Pesole, G., Saito, R., Wagner, L., Washio, T., Sakai, K., Okido, T., Furuno, M., Aono, H., Baldarelli, R., Barsh, G., Blake, J., Boffelli, D., Bojunga, N., Carninci, P., de Bonaldo, M. F., Brownstein, M. J., Bult, C., Fletcher, C., Fujita, M., Gariboldi, M., Gustincich, S., Hill, D., Hofmann, M., Hume, D. A., Kamiya, M., Lee, N. H., Lyons, P., Marchionni, L., Mashima, J., Mazzarelli, J., Mombauts, P., Nordone, P., Ring, B., Ringwald, M., Rodriguez, I., Sakamoto, N., Sasaki, H., Sato, K., Schonbach, C., Seya, T., Shibata, Y., Storch, K. F., Suzuki, H., Toyooka, K., Wang, K. H., Weitz, C., Whittaker, C., Wilming, L., Wyshaw-Boris, A., Yoshida, K., Hasegawa, Y., Kawaji, H., Kohtsuki, S. and Hayashizaki, Y.
 Functional annotation of a full-length mouse cDNA collection
 Nature 409 (6821), 685-690 (2001)
 21085660
 11217851
 5 The FANTOM Consortium and the RIKEN Genome Exploration Research Group Phase I & II Team.
 Analysis of the mouse transcriptome based on functional annotation of 60,770 full-length cDNAs
 Nature 420, 563-573 (2002)
 6 (bases 1 to 2927)
 Adachi, J., Aizawa, K., Akahira, S., Akimura, T., Araki, A., Aono, H., Arakawa, T., Bono, H., Carninci, P., Fukuda, S., Fukunishi, Y., Furuno, M., Hanagaki, T., Hara, A., Hayatsu, N., Hiramoto, K., Hirakawa, T., Hori, F., Imotani, K., Ishii, Y., Itoh, M., Izawa, M., Koyama, S., Kurihara, C., Kawai, J., Kojima, Y., Konno, H., Kouda, M., Nomura, K., Numazaki, R., Ohno, M., Okazaki, Y., Okido, T., Owa, C., Saito, H., Saito, R., Sakai, C., Sakai, K., Sano, H., Sasaki, D., Shibata, K., Shibata, Y., Shinagawa, A., Shiraki, T., Sogabe, Y.,


```

/clone_lib="NCI_CGAP_Lu30"
/notes="Organ: lung; Vector: pCMV-SPORT6; Site 1: NotI;
Site 2: SalI; transgenic model MNT-1, expression driven by
MTMV-LTR enhancer; Cloned unidirectionally. Primer: Oligo
dT. Library constructed by Life Technologies.
Investigator providing samples: Gilbert Smith, NIH"
BASE COUNT      280 a  222 g  222 c  222 g  230 t
ORIGIN

Query Match      39.8%; Score 505; DB 10; Length 954;
Best Local Similarity 85.3%; Pred. No. 3.2e-141;
Matches 598; Conservative 0; Mismatches 100; Indels 3; Gaps 3;

QY 493 CTTGCAAGTATAAATTTATTTTGCAGAAACGGGCCAGAAAGATCTATTATGTCGGC 552
DB 1  CGAGCTGTGTAATTAATCTGATCTTACAGAAACGGGCCAGAAAGATCTATTATGTCGGC 60

QY 553 TATTCACAGGGCACCAACCTATGAGCTTTATGCAATTTTCCACATGCCAGAGCTGGCTGAG 612
DB 61 TACTCACAGGGCACCAACCTATGAGCTTTATGCAATTTTCCACATGCCAGAGCTAGCTCAT 120

QY 613 AATATCAAAATGATTTTGTGCTTGTAGCACCATAGCCACTGTTAAGCATGCAAAAACCC 672
DB 121 AATATCAAAATGATTTTGTGCTTGTAGCACCATAGCCACTGTTAATATGCAAAAGTCT 180

QY 673 GGGACCAAAATTTTGTGCTGCGAGATGATCAAGGGATTTGTTGGCAAAAAGAA 732
DB 181 GGTACCAAAATTTTGTGCTGCGAGATGATCAAGGGATTTATTTGGCAGACAAGAG 240

QY 733 TTTCTGTATCAGACCAAGATTTCTCAGACACTTTGTTATTTACTTTTGGCCAGGTGATT 792
DB 241 TTTTATACCAAGTATGATTTTTCAGGAGCTTTTATTTTACC-TTGTGCGCCAGATGATT 299

QY 793 CTTGATCAGATTTGTATGATATATCATGTTACTTCTGGGTGGATTCAACACCAATATG 852
DB 300 CTCGACCAATCTGCAGACATCATCTTACTCTCTGGAGGTTTAAACACAAATATG 359

QY 853 AACATGAGCCGAGAGTATATGTCGCCACACACTCTTGTCTGGAAACATCTGTGCAAT 912
DB 360 AACATGAGCCGAGAGTATATGTCGCCACACACTCTTGTCTGGAAACATCTGTGCAAG 419

QY 913 ATTCTACATCTGGAGCAGGAGTGAATTTCTGGTGAATCTCGGGCATTTGACTGGGGAGT 972
DB 420 ATTCTACATCTGGAGCAGGAGTGAATTTCTGGTGAATCTCTGGGAACTTTGCTGGGAGT 479

QY 973 GAGACCAAAATCTGGAAAATGCAATTCAGCCAACTCTGTAAGGTACAGATCGAGAT 1032
DB 480 GAGACCAAAATCAGGAGAAATGTAATTCAGCCAACTCTTATAAGGTACAAAGTTCCAGAT 539

QY 1033 ATGACGCTCCCTACAGCAATGTGACAGAGGTCAGGACTGGCTTTCAATCCAGAGAC 1092
DB 540 ATGATGTTCCCAAGCAATGTGACAGGAGTCAAGACTGG-TTTCAAATCCAGATGAT 598

QY 1093 GTGAAAATGCTGCTCTCTGAGGTGACCACTCATCTACCATAGAATATTCCTGAATGG 1152
DB 599 GTGAAAATGCTGCTCTCTGAGGTGACCACTCATCTACCATAGAATATTCCTGAATGG 658

QY 1153 GCTCACGTGATTTTCACTCTGGGTTTGGATGCTCTCTCACCG 1193
DB 659 GGTATGTGGA-TTCATCTGGGGGGTGGATGCCCCCTCAGCG 698

RESULT 6
LOCUS      BF141567
DEFINITION 601788654P1 NCI_CGAP_Lu30 Mus musculus cdna clone IMAGE:4016413 5',
            mRNA sequence.
ACCESSION  BF141567
VERSION     BF141567.1 GI:10980607
KEYWORDS   EST.
SOURCE      Mus musculus (house mouse)
ORGANISM   Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;

Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
1 (Bases 1 to 1079)
NIH-WGC http://wgc.nci.nih.gov/
National Institutes of Health, Mammalian Gene Collection (MGC)
Unpublished
Contact: Robert Strausberg, Ph.D.
Email: rgs@bbs.fda.gov
Tissue Procurement: Gilbert Smith, Ph.D.
cDNA Library Preparation: Life Technologies, Inc.
cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)
DNA Sequencing by: InCytte Genomics, Inc.
Clone distribution: NCI-CGAP clone distribution information can be
found through the I.M.A.G.E. Consortium/LLNL at:
http://image.llnl.gov
Plate: LAM9263 row: 1 column: 14
High quality sequence stop: 663.
FEATURES
source
1..1079
/organism="Mus musculus"
/mol_type="mRNA"
/strain="C57BL/6J"
/db_xref="taxon:10090"
/clone="IMAGE:4016413"
/tissue_type="tumor, metastatic to mammary"
/lab_host="DH10B"
/clone_lib="NCI_CGAP_Lu30"
/notes="Organ: lung; Vector: pCMV-SPORT6; Site 1: NotI;
Site 2: SalI; transgenic model MNT-1, expression driven by
MTMV-LTR enhancer; Cloned unidirectionally. Primer: Oligo
dT. Library constructed by Life Technologies.
Investigator providing samples: Gilbert Smith, NIH"
BASE COUNT      312 a  267 c  263 g  237 t
ORIGIN

Query Match      39.2%; Score 497.6; DB 10; Length 1079;
Best Local Similarity 84.1%; Pred. No. 6.1e-139;
Matches 621; Conservative 0; Mismatches 109; Indels 8; Gaps 5;

QY 256 AAGAAGCAGGTTTCCAGGCTGTGGTTTACTCAGCATGGCTAGTGGAGTGGTAGC 315
DB 20 AAGAAGCAGGTTTCCAGGCTGTGGTTTACTCAGCATGGCTAGTGGAGTGGTAGC 79

QY 316 AACTGGATTTCCAACTGCCCAACAATAGCTGGGCTTCATTCTGGCAGATGCTGGTTT 375
DB 80 AACTGGATTTCCAACTGCCCAACAACAGCTGGGTTTATTCTGGCAGATGCTGGTTT 139

QY 376 GACGTGTGATGTTGGGAACAGCAGGGGAAACGCC-TGGTCTCGAAAACACAGACATCTC 434
DB 140 GATGTGTGATGTTGGGAACAGCAGGGGAAACACCGTGGTCTCGAAGCACAAGACCTCTC 199

QY 435 CATAGACCAAGATGAGTTCTGGGCTTTTCACTTATGATGATGGCTAGGTTTGACCTTCC 494
DB 200 CATAGATCAAGATGAGTTCTGGGCTTTTCACTTATGATGATGGCTAGGTTTGACCTTCC 259

QY 495 TCGAGTGATAAATCTTTATTTTGCAGAAAACGGGCCAGAAAAGATCTATTATGTCGGCTA 554
DB 260 AGCTGTGATAAATCTTTATTTACAGAAAACGGGCCAGAAAAGATCTATTATGTCGGCTA 319

QY 555 TTCACAGGGCACCACTGGCTTTTATGCAATTTTCCACCATGCCAGAGCTGGCTCAGAA 614
DB 320 CTCACAGGGCACCACTGGCTTTTATGCAATTTTCCACCATGCCAGAGCTAGCTCATAA 379

QY 615 AATCAAAATGATTTTGTCTTTAGCAACCATAGCCACTGTTTAAGCATGCAAAAAGCCCGG 674
DB 380 AATCAAAATGATTTTGTCTTTAGCCCTTAGCTCTATAGCCACTGTTTAATATGCAAAAGTCTCGG 439

QY 675 GACCAATTTTGTCTGCGCAGATATGATCAATGAGGATTTGTTGGCAAAAAGATTT 734
DB 440 TACCAATTTTGTCTGCGCAGATATGATCAATGAGGATTTTATTTGGCAGACAAGATTT 499

QY 735 TCTGTATCAGACCAAGATTTCTCAGACAACTTTGTTATTTTACCTTTGTGCCAGGTGATTCT 794
DB 500 TTTATACCAAGTATGATTTTTCAGGAGCTTTTATTTTACCTTTGTGCCAGATGATTCT 559

```


117	AGCTGTGGACCCAGAGACATTCTAGAAATATTAGTGAATCATCCAAATCATCAAGGCTATCC	176
Db		
140	AGCTGTGGATCCCTGAAGTCAACATGAATGACCGAGATAATCATCGCTGGGGATATCC	139
Qy		
177	CTGTGAGGAATATGAAGTCCCAACTGAAGATGGGTATATCTCTTTAAACAGGATTC	236
Db		
200	TGCTGAGGACACCTCGTCTCTGACAGGATGGCTATATTCTGAGTATTACCGAATCCC	259
Qy		
237	TCGAGGCTAGTGCAACCTTAAGAAGACAGGTTCCAGGCCCTGTGGTGTACTGCGAGCATGG	296
Db		
260	TCGTGGCGGGAAGAACATTTTGCCTAAGGTCCTCCAGACCAAGTGTGTGTATCTTCAGCACGG	319
Qy		
297	CCTAGTTGGGAGGTGCTAGCAACTGCATTTCCAACTGCCAACTCCAAATAGCCTGGCTTCAT	356
Db		
320	CTTTGCTGGCAGATCTCTAGTCTACTGGGTCAAAACAATTGACACAGCAGCTTTGGCTTCCT	379
Qy		
357	TCGTGGCAGATGCTGGTTTTGAGCTGTGGATGGGAAACAGCAGGGGAAACGCCTGGTCTCG	416
Db		
380	CCTGGCGGATGCTGGCTTTGATGTGTGGATGGGAAACAGCAGAGGAAACACCTGGTCTCT	439
Qy		
417	AAACAACAAGACACTCTCCATAGACCACAGATGAGTTCTGGGCTTTTCAGTTATGATGAGAT	476
Db		
440	GAAGCACAAGACTCTTTCAGTTCTCAGATGAGTTCTGGGCCCTTCAGTTTGTATGAGAT	499
Qy		
477	GGCTAGGTTTGACCTTCCTCGAGTGATAAATCTTTATTTGAGAAAAACGGGCCACGAAAAA	536
Db		
500	GGCAAAATATGACCTACCTGCTCAATATACTACATTTGAATAAAACTGGCCCAAGAACA	559
Qy		
537	GATCTATTATGTCGGCTATTACAGGGCACCAACATGGCTTTATTGCAATTTTCCACCAT	596
Db		
560	AATCTATTATGTGGGTCAATCTCAAGGCTGCACATAGTTTCATAGCTTTTTCACAGAT	619
Qy		
597	GCCAGAGCTGGCTCAGAAAAACAATAATATTTTGTCTTTAGCACCCATAGCCATGTGTAA	656
Db		
620	GCCTGAGTTGGCAAGAAGATTAAAAATCTTTTAGTCTTGCTCCGGTGTGTCTCTCAA	679
Qy		
657	GCATGCAAAAAGCCCGGGACCAAAATTTTGTGCTGCCAGATATGATGATCAAGGGATT	716
Db		
680	TTTTGTCTTCAGCGCCGCTACTTCAATTGGGACGCTTGCCACGCCCTCTTCTCAAGGAAT	739
Qy		
717	GTTTGGCAAAAAAGAAATTTCTGATACAGACCAGATTTCTCAGACAACTTGTTATTATCCT	776
Db		
740	GTTTGGACAGAAGCAATTTCTCCCTCAGAGTCCGATGCTGAAGTGCGCTGAGCATTACGT	799
Qy		
777	TTGTGGCCAGGTGATCTTGATCAGATTTGTATGTAATATCATGTACTTCTGGGTGGAAT	836
Db		
800	TTGCACCCACGTCATCATGAAGAGCTTTGGCCAAACGCTCTTCTTCCTGCTATGTGGCTT	859
Qy		
837	CAACACCAACAATATGAACATGAGCCGAGCAAGTGTATATGCTGCCACACACTTCTGCTGG	896
Db		
860	CNATGAGAAGAAATTTAAATATGCTTAGAGTGGATCTGTATACAACACACTGCCCTGCCGG	919
Qy		
897	AACATCTGTGAAAAATATTCTACATCTGAGCAGCGCAGTGAAATCTCGTGTAACTCCGGGC	956
Db		
920	AACTTCTGTGCAAAACATGTTTACACTGGGGTCAGGTTTTTAAATACCGGTAAAGCTTCAAGC	979
Qy		
957	ATTTGACTGGGGGAGTGAGACCAAAAACTGGAATAATGCAATCAGCCAACTCCTGTGAAG	1016
Db		
980	TTTTGACTGGGGAAGCAGTGAAGAAGAAATTTTTCATTACACCAAGATTTTCTCCCTC	1033
Qy		
1017	GTACAGAGTCAGATATGACGGTCCCTACAGCAATGTGACAGGAGGTCAAGACTGGCT	1076
Db		
1040	GTACAACATATAAAACATGCGTCTACCACTGCTTTGTGACGGGGGCCGGAATGGCT	1099


```

/db_xref="FANTOM DB:2310067K20"
/db_xref="MGI:1898392"
/db_xref="taxon:10090"
/cloned="2310067K20"
/sex="male"
/tissue_type="tongue"
/cloned_lib="RIKEN full-length enriched mouse cDNA library"
/dev_stage="adult"
53..1240
CDS
/notes="unnamed protein product; TRIACYLGLYCEROL LIPASE,
LINGUAL PRECURSOR (EC 3.1.1.3) (LINGUAL LIPASE) homolog
[Rattus norvegicus] (SWISSPROT|P04634, evidence: FASTV,
88.4%ID, 100%length, match=1184)
putative"
/codon_start=1
/protein_id="BAB26697.1"
/db_xref="MGI:12845298"
/db_xref="MGI:1914967"
/translation="MWLLVTVSLAFSGAHGLFGKLPKPNPNANNVSNQITWYGP
SEYEVVTDGVLGVIPYKQNSNENIKRPVAYLOHGLVASATNWTNLPNSLA
FILADAGYDMLGNSRGNTWPKNTVYSPDSVEFWAFSPDEMAKYDLPATIDFVOKT
GEKIHVSHSQTTIGFIATPSTNPAKAKIKRFYALAPVATVKYTESPPFKLSLPK
FLKLVFGNKMFPENYLDQFLGTEVCSRELLDLCSNALFICFGDKKNLNSRFDV
YLGNPAGTSTQDLFWAQLAKSGKLQAYNWSPLQNLHYNQKTPPYDYVSAMTVP
AYWNGGHDILADPDQVAMLLPKLPNLLYHKEILPYNHLDFIWMADAPQEVNEIVTM
AED"
BASE COUNT 396 a 290 c 286 g 378 t
ORIGIN
Query Match 36.4%; Score 461.8; DB 11; Length 1350;
Best Local Similarity 63.8%; Pred. No. 5e-128;
Matches 700; Conservative 0; Mismatches 397; Indels 0; Gaps 0;
QY 125 ACCCAGAGCATTCATGATATTAGTGAATCAATCAATCAAGGCTATCCTGTGAGG 184
DB 132 ACCTGAGCAACATGATGTTAGTCAGATGTAATCTTCTGGGATATCAAGTAGG 191
QY 185 AATATGAGTCGCAACTGAGATGAGTATATCTTCTGTTAACAGGATCTCTGAGGCC 244
DB 192 AATATGAGTTGTTACTGAGATGAGTATATCTTCTGGGATATCAAGTAGG 251
QY 245 TAGTGCACTAGACAGAGGTCAGGCTGAGTGTACTGAGGATGAGCTAGTGTG 304
DB 252 AGAAAAATTCGAGATATCGGCAAGAGACCTGTGCAATATTCAGCATGTTGGT 311
QY 305 GAGGTGCTAGCAACTGATTTCCAACTGCCCAACATAGCTGAGCTTCTTCTGCGCAG 364
DB 312 CATCAGCCCAAACTGGATTACAAATCTGCCAAACAGAGCTGCTTCTTCTAGCAG 371
QY 365 ATGCTGTTTGTGAGTGTGATGGGACAGAGGAGGAAACGCTGCTCGAAGACACA 424
DB 372 ATGCTGCTATGATGTGTGCTGGGAAACAGTCCAGGGAATACATGCTCCGGAATATG 431
QY 425 AGACACTCTCCATAGACAGATGAGTTCCTGGGCTTTCAGTTATGATGAGTGGCTAGT 484
DB 432 TATCTATTACACAGACTGATGATTCCTGGGCTTTCAGTTATGATGATGATGAT 491
QY 485 TTGACCTCTGCGATGATATACTTTATTTTTCAGAAAAACGGCCGAGGAAAGATCTATT 544
DB 492 ATGACCTTCCAGCCACCATAGATTCATTTGTACAGAAAACTGGACAGAGATACACT 551
QY 545 ATGTCGCTATTACAGGCGACCATGAGCTTTTATTTGATTTTCCACATGCGCAGAGC 604
DB 552 ATGTTGTCTACTCTCAGGCGACCATATCGTTTATTTGCTTTTCTTCACTCTGCTC 611
QY 605 TGGCTCAGAAAAATCAAAATGATTTTGTCTTTAGACCCATAGCCACTGTTAGCATGCAA 664
DB 612 TGGCTAAAAAATCAAGAGGTTTATGATGATGATGATGATGATGATGATGATGATGAT 671
QY 665 AAGCCCGGGGACCAATTTTGTGTCGCGCAGATGATGATGATGATGATGATGATGATGAT 724
DB 672 AAGTCCCTTTAAAAAGATTTCATTTATCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCT 731

```

```

QY 725 AAAAAGATTTCTGTATCAGACAGAGATTTCTCAGACAACTTTCTTATTTACCTTTTGGCC 784
DB 732 AAAAAATGTTTCATGCCCCCACTACTTAGATCAATTTCTTGGTAGGAGTGTGCTAC 791
QY 785 AGGTGATTTCTGTATCAGATTTGTAGTAATATCATGTCTTCTGGGTGATTTCAACCA 844
DB 792 GGGAGCTGTAGATCTTCTCTGAGCAACGCTTTATTCATCTTCTGTGATTTGACAAGA 851
QY 845 ACAAATGATGATGAGCGGAGCAAGTGTATATCTGCTCCCACTCTTCTGCTGGAACATCTG 904
DB 852 AAAAACTTAATGTGAGTGTGATGTGTATCTAGGCAATATCCAGGAGCAATCTA 911
QY 905 TGCAAAATATTTACACTGAGCAGCAGTGAATTTCTGTGAATCTCCGGGATTTGACT 964
DB 912 CTCAAGACCTTTTCCACTGGGACAGCTTCTTAATCTGGAAGCTTCAAGCTATAACT 971
QY 965 GGGGAGTGAGACAAATAATCTGAAAAATGATGATGATGATGATGATGATGATGATGATG 1024
DB 972 GGGGAGTCCATTTACAGAACATGTTACTACATCAGAAACGCTCCCTACTATGATG 1031
QY 1025 TCAGAGATATGACGTCCTTACAGCAATGTGGAGAGGAGTCCAGGACTGGCTTTCAAATC 1084
DB 1032 TGTGAGCAATGACCGTGCCAAATGTGAGTGTGGACGTTGCCATGACATCTCTGGCTGATC 1091
QY 1085 CAGAAGACGTGAAATATCTCTCTCTGAGGTGAGCAACCTCATCTACCATAGAAATATTC 1144
DB 1092 CCCAAGATGTGCAATGCTGCTTCCCAAACTCCCAACCTTCTGTACCATAGGAGATTC 1151
QY 1145 CTGAATGGCTCAGCTGATTTCTATCTGGGTTTGGATGCTCTCCTACCGTATGTACAATG 1204
DB 1152 TTCCCTACAATCACTGGATTTCTATCTGGCGATGATGGCGCTCAAGAGTTTACATG 1211
QY 1205 AAATCATCCATCTGATG 1221
DB 1212 AGATAGTTACCATGATG 1228

RESULT 10
AK010116 1347 bp mRNA linear HTC 05-DEC-2002
LOCUS Mus musculus adult male tongue cDNA, RIKEN full-length enriched
DEFINITION library, clone:2310068005 product:TRIACYLGLYCEROL LIPASE, LINGUAL
PRECUSOR (EC 3.1.1.3) (LINGUAL LIPASE) homolog (Rattus
norvegicus), full insert sequence.
ACCESSION AK010116
VERSION AK010116.1 GI:12845334
KEYWORDS HTC; CAP trapper.
SOURCE Mus musculus (house mouse)
ORGANISM Mus musculus
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
1
Carninci, P. and Hayashizaki, Y.
High-efficiency full-length cDNA cloning
Meth. Enzymol. 303, 19-44 (1999)
99279253
10349636
2
Carninci, P. and Hayashizaki, Y.
Itoh, M., Konno, H., Okazaki, Y., Muramatsu, M. and Hayashizaki, Y.
Normalization and subtraction of cap-trapper-selected cDNAs to
prepare full-length cDNA libraries for rapid discovery of new genes
Genome Res. 10 (10), 1617-1630 (2000)
20499374
11042159
3
Shibata, K., Itoh, M., Aizawa, K., Nagasaka, S., Sasaki, N., Carninci, P.,
Konno, H., Akiyama, J., Nishi, K., Kitsuai, T., Tashiro, H., Itoh, M.,
Sumi, N., Ishii, Y., Nakamura, S., Hazama, M., Nishino, T., Harada, A.,
Yanamoto, R., Matsumoto, H., Sakaguchi, S., Ikegami, T., Kashiwagi, K.,
Fujiwara, S., Inoue, K., Togawa, Y., Izawa, M., Ohara, E., Watahiki, M.,
Yoneda, Y., Ishikawa, T., Ozawa, K., Tanaka, T., Matsuura, S., Kawai, J.,
Okazaki, Y., Muramatsu, M., Inoue, Y., Kira, A. and Hayashizaki, Y.

```



```

source
1. .1349
/organism="Mus musculus"
/mol_type="mRNA"
/strain="C57BL/6J"
/db_xref="PANTOM DB:2310011G18"
/db_xref="MGI:1895751"
/db_xref="taxon:10090"
/clone="2310011G18"
/sex="male"
/tissue_type="tongue"
/clone_lib="RIKEN full-length enriched mouse cDNA library"
/dev_stage="adult"
53. 1240
/note="unnamed protein product; TRIACYLGLYCEROL LIPASE, LINGUAL PRECURSOR (EC 3.1.1.3) (LINGUAL LIPASE) homolog [Rattus norvegicus] (SWISSPROT|P04634, evidence: FASTA, 88.4%ID, 100%length, match=1184) putative"
/codon_start=1
/protein_id="BAB26201.1"
/db_xref="GI:12844012"
/db_xref="MGI:1914967"
/translation="MWLLLVSVLSAFGGHGLFKLGPKNPEANMNVSMITWGVPP SEEVVTDYDILGVNRPYKKNSENIGRPVAYLQHLIASATNMTLNENSLA FILADAGVWLGSRNWTGRKNVYSPVFAFSDPMAYKIDLPIDTIVQKT GQEKHYVGHSGQTIGFIAFTNPALAKIKRYFALAPVATVYTESPEPKISLIPK FLKVIYFNKMFPHNYLDQFLGTVEVCSRELLDLCSNALFIFCGPKNLSVRFQV YLGNPACTSDQLPFWAQLAKSKLOAYNWGSPLQNMHLNOKTPPYVDVSATVPI AVWNGHIDILADPDQVAMLLPKLPNLLYHKELLPNHLDPFWANDAPQEVNIVTM AED"
BASE COUNT      395 a 290 c 284 g 380 t
ORIGIN

Query Match      36.3%; Score 460.2; DB 11; Length 1349;
Best Local Similarity 63.7%; Pred No. 1.5e-127;
Matches 699; Conservative 0; Mismatches 398; Indels 0; Gaps 0;

QY 125 ACCGAGAGCATTCATGAATATTAGTGAATCATCCAAATCAAGGCTATCCCTGTGAGG 184
DB 132 ACCCTGAAGCAACATGAATGTTAGTCAGATGATAACTTACTGGGGATATCCAAAGTCAGG 191
QY 185 AATATGAAGTCGCACTGAGTGGTATATCCCTTCTGTGTAACAGGATTCCTCGAGGCC 244
DB 192 AATATGAAGTTGTTACTGAAGATGCTACATTCCTGGGGGCTCTAGAGATTCCTTATGGGA 251
QY 245 TAGTGCAACCTAAGAGACAGGTTCCAGGCTGTGTGTACTGCGAGCATGCGCTAGTTG 304
DB 252 AGAAAAATTCGAGAATATCGCAAGACACCTGTGGCATAATTCGAGCATGTTTGTATTG 311
QY 305 GAGGTGCTAGCACTGGATTTCCAACTGCGCAACATAGCTGGGCTTCATTCCTGGCAG 364
DB 312 CATCAGCCAACTGGGATTAACAATCTGCCAAACACAGCGCTGGCCTTCATTCCTAGCAG 371
QY 365 ATGCTGGTTTTCAGCTGTGGTGGGGAACAGCAGGGGAAACGCTGGTCTCGAAAAACA 424
DB 372 ATGCTGGCTATGATGTGTGGTGGGGAACAGCTCGAGGAATACATGTCCCGGAAAAATG 431
QY 425 AGACACTCTCCATAGACCAAGATAGTTCTGGGCTTTTCAGTTATGATGATGCTAGGT 484
DB 432 TATACTATTCCACGACTCAGTTGAATTCCTGGGCTTTTCAGCTTTGATGAATGCTAAAT 491
QY 485 TTGACCTCTCTGCTGATGATACTTTATTTGCAAAAACGGCGCAGGAAGATCTATT 544
DB 492 ATGACTTTCAGCCACCTAGACTTCATTTGACAGAAACCTGGACAGAGAGATACACT 551
QY 545 ATGTCGGCTATTACAGGGCCACCACTGGGCTTTTATGCTATTTTCCACCATGCCAGGC 604
DB 552 ATGTTGGTCACTCTCAGGSCACCACTATCGGTTTTATTTGCTTTTCTTACCAATCCTGTC 611
QY 605 TGGCTCAGAAATCAAAATGATTTTGTCTTAGCACCCCATAGCCACTGTTAAGCATGCAA 664
DB 612 TGGCTAAAAAATCAAGAGGTTTTATGATGATGCTTCCAGTTGCTACTGTGAAGTATACAG 671

```

AK010139 1350 bp mRNA linear HTC 05-DEC-2002
Mus musculus adult male tongue cDNA, RIKEN full-length enriched library, clone:2310069P19 product:TRIACYLGLYCEROL LIPASE, LINGUAL PRECURSOR (EC 3.1.1.3) (LINGUAL LIPASE) homolog [Rattus norvegicus], full insert sequence.

AK010139 1 GI:12845371
HTC; CAP trapper.
Mus musculus (house mouse)
Mus musculus
Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.

1 Carninci, P. and Hayashizaki, Y.
High-efficiency full-length cDNA cloning
Meth. Enzymol. 303, 19-44 (1999)
99279253
10349636
2
3 Carninci, P., Shibata, Y., Hayatsu, N., Sugahara, Y., Shibata, K., Itoh, M., Konno, H., Okazaki, F., Muramatsu, M. and Hayashizaki, Y.
Normalization and subtraction of cap-trapper-selected cDNAs to prepare full-length cDNA libraries for rapid discovery of new genes
Genome Res. 10 (10), 1617-1630 (2000)
20499374
11042159
3
Shibata, K., Itoh, M., Aizawa, K., Nagaoka, S., Sasaki, N., Carninci, P., Konno, H., Akiyama, J., Nishi, K., Kitsumai, T., Tashiro, H., Itoh, M., Sumi, N., Ishii, Y., Nakamura, S., Hazama, M., Nishine, T., Harada, A.,

Db 612 TGGCTAAAAAATCAAGAGGTTTTATGCTTAGCTCCAAATTTGCTACTGTGAAGTATACAG 671

Qy 665 AAAGCCCGGAGCAAAATTTTGTCTGCGCAGATATGATCAATCAAGGATTTTGGCA 724

Db 672 AAATGCTTTTAAAGAGATTTTCACTTATTCCTAAGTTTCTTCAAGGTGATATTTGGTA 731

Qy 725 AAAAAGAAATTTCTGTATCAGACAGATTTCTTCAGACAACTTTGTTATTTACTTTTGGCC 784

Db 732 AAAAAATGTTTATGCCCAACACTTACTAGATCAATTTCTTGTACGGAAGTGTGCTCAC 791

Qy 785 AGGTGATTTTATCAGATTTTGTAGTATATATCTTCTGCTGGTGGATTTCAACACCA 844

Db 792 GGGAGCTGCTAGATTTCTTCTGCAAGCAAGCTTTTATTTATCTTCTGTGGATTTGACAAGA 851

Qy 845 ACAATATGAACATCAGCGCAGCAAGTGTATATGCTGCCACACTTTGCTGGCAACATCTG 904

Db 852 AAARCTTAAATGTGATCGCTTTGATGTATCTAGGGCATATCCAGACAGGACATCTA 911

Qy 905 TGCATAATTTTCACTGTGAGCAGGAGTGAATTTCTGTGAATCTCGGGCATTTGACT 964

Db 912 CTCAAGACCTTTTCCACTGGGCACAGCTTGTCTAAATCTGGAGCTTCAAGCCTATAACT 971

Qy 965 GGGGAGGTGACCAAAATCTGGAATTCGAATTCGAATTCGAATTCGAATTCGAATTCGAATTC 1024

Db 972 GGGGAGGTGACCAAAATCTGGAATTCGAATTCGAATTCGAATTCGAATTCGAATTCGAATTC 1031

Qy 1025 TCAGAGATATCAGCTCCCTACAGCAATGTGGACAGGAGGTGAGGACTTTCGAATTC 1084

Db 1032 TGTGAGCCTACAGCTGCAATTCAGTGTGGACAGGAGGTGAGGACTTTCGAATTC 1091

Qy 1085 CAGAGAGCTGAAATGCTGCTCTGAGGTGAGGAGGTGAGGACTTTCGAATTC 1144

Db 1092 CCCAGAGGTGCAATGCTGCTTCCCAACTTCCCAACTTCCCAACTTCCCAACTTCCCAACTT 1151

Qy 1145 CTGAATGGCTCAGCTGATTTCTCTGGGTTTGGATGCTTCCCTCAGGATGTACAAATG 1204

Db 1152 TTCCCTACATCACTGGACTTCTCTGGGCGATGGATGGCTCAGAGGTTTACATG 1211

Qy 1205 AAATCATCCTCTGATG 1221

Db 1212 AGATAGTTACCATGATG 1228

RESULT 13

AK010124

LOCUS

DEFINITION Mus musculus adult male tongue cDNA, RIKEN full-length enriched library, clone:2310069D09 product:TRIACYLGLYCEROL LIPASE, LINGUAL PRBCURSOR (EC 3.1.1.3) (LINGUAL LIPASE) homolog [Rattus norvegicus], full insert sequence.

ACCESSION AK010124

VERSION AK010124.1

KEYWORDS GI:12845346

SOURCE HTG; CAP trapper.

ORGANISM Mus musculus (house mouse)

REFERENCE

1 Carninci, P., Shibata, Y., Hayatsu, N., Sugahara, Y., Shibata, K., Itoh, M., Kono, H., Okazaki, Y., Muramatsu, M., and Hayashizaki, Y. Normalization and subtraction of cap-trapper-selected cDNAs to prepare full-length cDNA libraries for rapid discovery of new genes Genome Res. 10 (10), 1617-1630 (2000)

2 20499374

3 11042159

4 10349636

5 10349636

6 10349636

7 10349636

8 10349636

9 10349636

10 10349636

11 10349636

12 10349636

13 10349636

14 10349636

15 10349636

16 10349636

17 10349636

18 10349636

19 10349636

20 10349636

21 10349636

22 10349636

23 10349636

24 10349636

25 10349636

26 10349636

27 10349636

28 10349636

29 10349636

30 10349636

31 10349636

32 10349636

33 10349636

34 10349636

35 10349636

36 10349636

37 10349636

38 10349636

39 10349636

40 10349636

41 10349636

42 10349636

43 10349636

44 10349636

45 10349636

46 10349636

47 10349636

48 10349636

49 10349636

50 10349636

51 10349636

52 10349636

53 10349636

54 10349636

55 10349636

56 10349636

57 10349636

58 10349636

59 10349636

60 10349636

61 10349636

62 10349636

63 10349636

64 10349636

65 10349636

66 10349636

67 10349636

68 10349636

69 10349636

70 10349636

71 10349636

72 10349636

73 10349636

74 10349636

75 10349636

76 10349636

77 10349636

78 10349636

79 10349636

80 10349636

81 10349636

82 10349636

83 10349636

84 10349636

85 10349636

86 10349636

87 10349636

88 10349636

89 10349636

90 10349636

91 10349636

92 10349636

93 10349636

94 10349636

95 10349636

96 10349636

97 10349636

98 10349636

99 10349636

100 10349636

101 10349636

102 10349636

103 10349636

104 10349636

105 10349636

106 10349636

107 10349636

108 10349636

109 10349636

110 10349636

111 10349636

112 10349636

113 10349636

114 10349636

115 10349636

116 10349636

117 10349636

118 10349636

119 10349636

120 10349636

121 10349636

122 10349636

123 10349636

124 10349636

125 10349636

126 10349636

127 10349636

128 10349636

129 10349636

130 10349636

131 10349636

132 10349636

133 10349636

134 10349636

135 10349636

136 10349636

137 10349636

138 10349636

139 10349636

140 10349636

141 10349636

142 10349636

143 10349636

144 10349636

145 10349636

146 10349636

147 10349636

148 10349636

149 10349636

150 10349636

151 10349636

152 10349636

153 10349636

154 10349636

155 10349636

156 10349636

157 10349636

158 10349636

159 10349636

160 10349636

161 10349636

162 10349636

163 10349636

164 10349636

165 10349636

166 10349636

167 10349636

168 10349636

169 10349636

170 10349636

171 10349636

172 10349636

173 10349636

174 10349636

175 10349636

176 10349636

177 10349636

178 10349636

179 10349636

180 10349636

181 10349636

182 10349636

183 10349636

184 10349636

185 10349636

186 10349636

187 10349636

188 10349636

189 10349636

190 10349636

191 10349636

192 10349636

193 10349636

194 10349636

195 10349636

196 10349636

197 10349636

198 10349636

199 10349636

200 10349636

201 10349636

202 10349636

203 10349636

204 10349636

205 10349636

206 10349636

207 10349636

208 10349636

209 10349636

210 10349636

211 10349636

212 10349636

213 10349636

214 10349636

215 10349636

216 10349636

217 10349636

218 10349636

219 10349636

220 10349636

221 10349636

222 10349636

223 10349636

224 10349636

225 10349636

226 10349636

227 10349636

228 10349636

229 10349636

230 10349636

231 10349636

232 10349636

233 10349636

234 10349636

235 10349636

236 10349636

237 10349636

238 10349636

239 10349636

240 10349636

241 10349636

242 10349636

243 10349636

244 10349636

245 10349636

246 10349636

247 10349636

248 10349636

249 10349636

250 10349636

251 10349636

252 10349636

253 10349636

254 10349636

255 10349636

256 10349636

257 10349636

258 10349636

259 10349636

260 10349636

261 10349636

262 10349636

263 10349636

264 10349636

265 10349636

266 10349636

267 10349636

268 10349636

269 10349636

270 10349636

271 10349636

272 10349636

273 10349636

274 10349636

275 10349636

276 10349636

277 10349636

278 10349636

279 10349636

280 10349636

281 10349636

282 10349636

283 10349636

284 10349636

285 10349636

286 10349636

287 10349636

288 10349636

289 10349636

290 10349636

291 10349636

292 10349636

293 10349636

294 10349636

295 10349636

296 10349636

297 10349636

298 10349636

299 10349636

300 10349636

301 10349636

302 10349636

303 10349636

304 10349636

305 10349636

306 10349636

307 10349636

308 10349636

309 10349636

310 10349636

311 10349636

312 10349636

313 10349636

314 10349636

315 10349636

316 10349636

317 10349636

318 10349636

319 10349636

320 10349636

321 10349636

322 10349636

323 10349636

324 10349636

325 10349636

326 10349636

327 10349636

328 10349636

329 10349636

330 10349636

331 10349636

332 10349636

333 10349636

334 10349636

335 10349636

336 10349636

337 10349636

338 10349636

339 10349636

340 10349636

341 10349636

342 10349636

343 10349636

344 10349636

345 10349636

346 10349636

347 10349636

348 10349636

349 10349636

350 10349636

351 10349636

352 10349636

353 10349636

354 10349636

355 10349636

356 10349636

357 10349636

358 10349636

359 10349636

360 10349636

361 10349636

362 10349636

363 10349636

364 10349636

365 10349636

366 10349636

367 10349636

368 10349636

369 10349636

370 10349636

371 10349636

372 10349636

373 10349636

374 10349636

375 10349636

376 10349636

377 10349636

378 10349636

379 10349636

380 10349636

381 10349636

382 10349636

383 10349636

384 10349636

385 10349636

386 10349636

387 10349636

388 10349636

389 10349636

390 10349636

391 10349636

392 10349636

393 10349636

394 10349636

395 10349636

396 10349636

397 10349636

398 10349636

399 10349636

400 10349636

401 10349636

402 10349636

403 10349636

404 10349636

405 10349636

406 10349636

407 10349636

408 10349636

409 10349636

410 10349636

411 10349636

412 10349636

413 10349636

414 10349636

415 10349636

416 10349636

417 10349636

418 10349636

419 10349636

420 10349636

421 10349636

422 10349636

423 10349636

424 10349636

425 10349636

426 10349636

427 10349636

428 10349636

429 10349636

430 10349636

431 10349636

432 10349636

433 10349636

434 10349636

435 10349636

436 10349636

437 10349636

438 10349636

439 10349636

440 10349636

441 10349636

442 10349636

443 10349636

444 10349636

445 10349636

446 10349636

447 10349636

448 10349636

449 10349636

450 10349636

451 10349636

452 10349636

453 10349636

454 10349636

455 10349636

456 10349636

457 10349636

458 10349636

459 10349636

460 10349636

461 10349636

462 10349636

463 10349636

464 10349636

465 10349636

466 10349636

467 10349636

468 10349636

469 10349636

470 10349636

471 10349636

472 10349636

473 10349636

474 10349636

475 10349636

476 10349636

477 10349636

478 10349636

479 10349636

480 10349636

481 10349636

482 10349636

483 10349636

484 10349636

485 10349636

486 10349636

487 10349636

488 10349636

489 10349636

490 10349636

491 10349636

492 10349636

493 10349636

494 10349636

495 10349636

496 10349636

497 10349636

498 10349636

499 10349636

500 10349636

501 10349636

502 10349636

503 10349636

504 10349636

505 10349636

506 10349636

507 10349636

508 10349636

509 10349636

510 10349636

511 10349636

512 10349636

513 10349636

514 10349636

515 10349636

516 10349636

517 10349636

518 10349636

519 10349636

520 10349636

521 10349636

522 10349636

523 10349636

524 10349636

525 10349636

526 10349636

527 10349636

528 10349636

529 10349636

530 10349636

531 10349636

532 10349636

533 10349636

534 10349636

535 10349636

536 10349636

537 10349636

538 10349636

539 10349636

540 10349636

541 10349636

542 10349636

543 10349636

544 10349636

545 10349636

546 10349636

547 10349636

548 10349636

549 10349636

550 10349636

551 10349636

552 10349636

553 10349636

554 10349636

555 10349636

556 10349636

557 10349636

558 10349636

559 10349636

560 10349636

561 10349636

562 10349636

563 10349636

564 10349636

565 10349636

566 10349636

567 10349636

568 10349636

569 10349636

570 10349636

571 10349636

572 10349636

573 10349636

574 10349636

575 10349636

576 10349636

577 10349636

578 10349636

579 10349636

580 10349636

581 10349636

582 10349636

583 10349636

584 10349636

585 10349636

586 10349636

587 10349636

588 10349636

589 10349636

590 10349636

591 10349636

592 10349636

593 10349636

594 10349636

595 10349636

596 10349636

597 10349636

598 10349636

599 10349636

600 10349636

601 10349636

602 10349636

603 10349636

604 10349636

605 10349636

606 10349636

607 10349636

608 10349636

609 10349636

610 10349636

611 10349636

612 10349636

613 10349636

614 10349636

615 10349636

616 103496

REFERENCE AUTHORS

3. Shibata, K., Itoh, M., Aizawa, K., Nagaoka, S., Sasaki, N., Carninci, P., Konno, H., Akiyama, J., Nishi, K., Kitsuai, T., Tashiro, H., Itoh, M., Sumi, N., Ishii, Y., Nakamura, S., Hazama, M., Nishize, T., Harada, A., Yamamoto, R., Matsumoto, H., Sakaguchi, S., Ikegami, T., Kashiwagi, K., Fujiwara, S., Inoue, K., Togawa, Y., Izawa, M., Ohara, E., Watahiki, M., Yoneda, Y., Ishikawa, T., Ogawa, K., Tanaka, T., Matsura, S., Kawai, J., Okazaki, Y., Muramatsu, M., Inoue, Y., Kira, A. and Hayashizaki, Y. RIKEN integrated sequence analysis (RISA) system - 384-format sequencing pipeline with 384 multicapillary sequencer (Genome Res. 10 (11), 1757-1771 (2000))

TITLE	JOURNAL	MEDLINE	PUBMED	REFERENCE	AUTHOR
1. The effect of the... 2. The effect of the... 3. The effect of the... 4. The effect of the... 5. The effect of the... 6. The effect of the... 7. The effect of the... 8. The effect of the... 9. The effect of the... 10. The effect of the...	1. The effect of the... 2. The effect of the... 3. The effect of the... 4. The effect of the... 5. The effect of the... 6. The effect of the... 7. The effect of the... 8. The effect of the... 9. The effect of the... 10. The effect of the...	1. The effect of the... 2. The effect of the... 3. The effect of the... 4. The effect of the... 5. The effect of the... 6. The effect of the... 7. The effect of the... 8. The effect of the... 9. The effect of the... 10. The effect of the...	1. The effect of the... 2. The effect of the... 3. The effect of the... 4. The effect of the... 5. The effect of the... 6. The effect of the... 7. The effect of the... 8. The effect of the... 9. The effect of the... 10. The effect of the...	1. The effect of the... 2. The effect of the... 3. The effect of the... 4. The effect of the... 5. The effect of the... 6. The effect of the... 7. The effect of the... 8. The effect of the... 9. The effect of the... 10. The effect of the...	1. The effect of the... 2. The effect of the... 3. The effect of the... 4. The effect of the... 5. The effect of the... 6. The effect of the... 7. The effect of the... 8. The effect of the... 9. The effect of the... 10. The effect of the...

Kawai, J., Shinagawa, A., Shibata, K., Yoshino, M., Itoh, M., Ishii, Y., Arakawa, I., Hara, A., Fukunishi, Y., Konno, H., Adachi, J., Fukuda, S., Aizawa, K., Iwaza, M., Nishi, K., Kiyosawa, H., Kondo, S., Yamakura, I., Saito, T., Okazaki, Y., Gojobori, T., Bono, H., Kasukawa, T., Saito, R., Kadota, K., Matsuda, H., Ashburner, M., Batalov, S., Casavant, T., Fleischmann, W., Gaasterland, T., Gissi, C., King, B., Kochiwa, H., Kuehl, P., Lewis, S., Matsumoto, Y., Nikaide, I., Pesole, G., Quackenbush, J., Schriml, L. M., Stauble, J. F., Suzuki, R., Tomita, M., Wagner, L., Washio, T., Sakai, K., Okido, T., Furuno, M., Aono, H., Baldarelli, R., Barsh, G., Blake, J., Boffelli, D., Bojunga, N., Carninci, P., de Bonaldo, M. F., Brownstein, M. J., Bult, C., Fletcher, C., Fujita, A. A., Kamiya, M., Lee, N. H., Lyons, P., Hofmann, M., Hume, D. A., Gariboldi, M., Guscicich, S., Hill, D., Marchionni, L., Mashima, Y., Mazzarelli, J., Mombaerts, P., Nordone, P., Ring, B., Ringwald, M., Rodrigues, I., Sakamoto, N., Sasaki, H., Sato, K., Schonbach, C., Seya, T., Shibata, Y., Storch, K. F., Suzuki, H., Toyooka, K., Wang, K. H., Weitz, C., Whittaker, C., Wilming, L., Wynshaw-Boris, A., Yoshida, K., Hasegawa, Y., Kawaji, H., Kohtsuki, S., and Havashizaki, Y.

TITLE
JOURNAL
MEDLINE
PUBMED
REFERENCE
AUTHORS

11217051
5
The FANTOM Consortium and the RIKEN Genome Exploration Research
Group Phase I & II Team.
Analysis of the mouse transcriptome based on functional annotation
of 60,770 full-length cDNAs
Nature 420, 563-573 (2002)
6 (bases 1 to 1343)
6

Db 549 ATGTTGGTCACTCTCAGGCGACCACTATCGGTTTATTGCTTTTCTTACCAATCTGCTC 608
Qy 605 TGCTCAGAAATCAAAATGTTTCTTTAGCAGCCATAGCCACTGTTAGCATGCA 664
Db 609 TGCTCAAAAATCAAGAGGTTTATGTCATTAGCTCCAGTCTGCTAGTGAATATACAG 668
Qy 665 AAAGCCCGGCGCAAAATTTTCTGCTCCAGATATGATCATCAAGGATTTGTTGGCA 724
Db 669 AAAGTCCCTTAAAGAGTTTCACTTATTCCTTAAGTTTCTTCTCAAGGATATTTGTA 728
Qy 725 AAAAAGATTTCTGATACAGACAGATTTCTCAGACAACTGTTATTTACCTTTGTCGCC 784
Db 729 ACAAATGTTTCATGCCCCCACTACTTAGATCAATTTCTTGTACGGAAGTGTGCTCAC 788
Qy 785 AGTGATTTCTGATCAGATTTGATGTAATATCATGTTACTTCTGGGTGATTCACACCA 844
Db 789 GGGAGCTGTGATGTTCTCTCGAGCAAGCTTTATTCATCTCTGTTGATTTGACAGA 848
Qy 845 ACAATATGAATGAGCGGAGCAAGTGTATGCTGCCCCACACTCTTTGCTGGAACATCTG 904
Db 849 AAAAATTAATGAGTGGCTTTGATGTTATCTAGGGCATATCCAGCAGGAAACATCTA 908
Qy 905 TCGAATATTTACACTGAGCGAGCAGCACTGATTTCTGTGAATCTCGGCACTTTGACT 964
Db 909 CTCAGACCTTTTCCACTGGGCGAGCTTCTTAATTTGGGAAGCTTCAAGCCTTAAT 968
Qy 965 GGGGAGTGTAGACCAAAATCTGGAATAATGCAATCAGCAACTCTCTGTAAGTACAGAG 1024
Db 969 GGGGAAGTCCATTTACAGACATGTTACATACATACAGAAACGCTCTCTACTATGATG 1028
Qy 1025 TCAGATATGACGGTCCCTACAGCATGTTGACAGAGTGTGAGGCTGAGTCTTCAATC 1084
Db 1029 TGTGAGCATGACCGTCCCAATTTGATGTTGAGTGTGAGGCTGAGTCTTGGCTGATC 1088
Qy 1085 CAGAAGACGTGAAATGCTCTCTCAGGTGACCAACCTCATCTACCAATGAAATATTC 1144
Db 1089 CCCAAGATGTGCAATGCTGCTCCCAACCTCCCAACCTCTGTACCAATGAGATTC 1148
Qy 1145 CTGAATGGGCTCAGTGTGATTTGATCTGGGGTTGGATGCTCTCCACCGTATGTAATG 1204
Db 1149 TTCCCTCAATCAGCTGAGTTCATCTGGGCGATGATGCGCTCAAGAGGTTTCAATG 1208
Qy 1205 AAATCATCCATCTGATG 1221
Db 1209 AGATAGTTACCATGATG 1225

RESULT 15
AC009413
LOCUS
DEFINITION
Mus musculus adult male tongue cDNA, RIKEN full-length enriched library, clone:2310020222 product:TRACVGLYCEPOT LIPASE, LINGUAL PRECURSOR (EC 3.1.1.3) (LINGUAL LIPASE) homolog (Rattus norvegicus), full insert sequence.
AC009413
VERSION
KEYWORDS
SOURCE
ORGANISM
Mus musculus (house mouse)
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
1
Carninci, P., Shibata, Y., Hayatsu, N., Sugahara, Y., Shibata, K., Itoh, M., Konno, H., Okazaki, Y., Muramatsu, M. and Hayashizaki, Y. Normalization and subcloning of cap-trapper-selected cDNAs to prepare full-length cDNA libraries for rapid discovery of new genes Genome Res. 10 (10), 1617-1630 (2000)
2
10349636
REFERENCE
AUTHORS
TITLE
JOURNAL
MEDLINE
PUBMED
REFERENCE
AUTHORS
TITLE
JOURNAL

20499374
11042159
3
Shibata, K., Itoh, M., Aizawa, K., Nagaoka, S., Sasaki, N., Carninci, P., Konno, H., Akiyama, J., Nishi, K., Katsunari, T., Tashiro, H., Itoh, M., Sumi, Y., Ishii, Y., Nakamura, S., Hazama, M., Nishine, F., Harada, A., Yamamoto, R., Matsumoto, H., Sakaguchi, S., Ikegami, T., Kashiwagi, K., Fujiwara, S., Inoue, K., Togawa, Y., Izawa, M., Ohara, E., Watanabe, M., Yoneda, Y., Ishikawa, T., Ozawa, K., Tanaka, T., Matsura, S., Kawai, J., Okazaki, Y., Muramatsu, M., Inoue, Y., Kira, A. and Hayashizaki, Y.
RIKEN integrated sequence analysis (RISA) system--384-format sequencing pipeline with 384 multiplexed capillary sequencer
Genome Res. 10 (11), 1757-1771 (2000)
20530913
11076861
4
Kawai, J., Shinagawa, A., Shibata, K., Yoshino, M., Itoh, M., Ishii, Y., Arakawa, T., Hara, A., Fukunishi, Y., Konno, H., Adachi, J., Fukuda, S., Aizawa, K., Izawa, M., Nishi, K., Kiyosawa, H., Kondo, S., Yamana, I., Saito, T., Okazaki, Y., Gojohori, T., Bono, H., Kasukawa, T., Saito, R., Kadota, K., Matsuda, H., Ashburner, M., Batalov, S., Casavant, T., Fleischmann, W., Gaasterland, T., Gissi, C., King, B., Kochiwa, H., Kuehl, P., Lewis, S., Matsuo, Y., Nikaide, I., Pesole, G., Quackenbush, J., Schriml, L. M., Stauber, F., Suzuki, R., Tomita, M., Wagner, L., Washio, T., Sakai, K., Okido, T., Furuno, M., Aono, H., Baldarelli, R., Barsh, G., Blake, J., Boffelli, D., Bojunga, N., Carninci, P., de Bonaldo, M. F., Brownstein, M. J., Bult, C., Fletcher, C., Fujita, M., Gariboldi, M., Gustincich, S., Hill, D., Hofmann, M., Hume, D. A., Kamiya, M., Lee, N. H., Lyons, P., Marchionni, L., Mashima, J., Mazzarelli, J., Mombauts, P., Nordone, P., Ring, B., Ringwald, M., Rodriguez, I., Sakamoto, N., Sasaki, H., Sato, K., Schonbach, C., Seva, T., Shibata, Y., Storch, K. P., Suzuki, H., Toyooka, K., Wang, K. H., Weitz, C., Wittaker, C., Wilming, L., Wynshaw-Boris, A., Yoshida, K., Hasegawa, Y., Kawaji, H., Kohetsuki, S. and Hayashizaki, Y.
Functional annotation of a full-length mouse cDNA collection
Nature 409 (6821), 685-690 (2001)
21085660
11217851
5
The PANTOM Consortium and the RIKEN Genome Exploration Research Group Phase I & II Team.
Analysis of the mouse transcriptome based on functional annotation of 60,770 full-length cDNAs
Nature 420, 563-573 (2002)
6 (bases 1 to 1345)
Adachi, J., Aizawa, K., Akahira, S., Akimura, T., Arai, A., Aono, H., Arakawa, T., Bono, H., Carninci, P., Fukuda, S., Fukunishi, Y., Furuno, M., Hanagaki, T., Hara, A., Hayatsu, N., Hiramoto, K., Hirao, T., Hori, F., Imotani, K., Ishii, Y., Itoh, M., Izawa, M., Kasukawa, T., Kato, H., Kawai, J., Kojima, Y., Konno, H., Kouda, M., Koya, S., Kurihara, C., Matsuyama, T., Miyazaki, A., Nishi, K., Nomura, K., Numazaki, R., Ohno, M., Okazaki, Y., Okido, T., Owa, C., Saito, H., Saito, R., Sakai, C., Sakai, K., Sano, H., Sasaki, D., Shibata, K., Shibata, Y., Shingawa, A., Shiraki, T., Sugabe, Y., Suzuki, H., Tagami, M., Tagawa, A., Takahashi, F., Tanaka, I., Tejima, Y., Toya, T., Yamamura, T., Yasunishi, A., Yoshida, K., Yoshino, M., Muramatsu, M. and Hayashizaki, Y.
Direct Submission
Submitted (10-JUL-2000) Yoshihide Hayashizaki, The Institute of Physical and Chemical Research (RIKEN) Laboratory for Genome Exploration and Research Group, RIKEN Genomic Sciences Center (GSC), RIKEN Yokohama Institute, 1-7-22 Suehiro-cho, Tsurumi-ku, Yokohama, Kanagawa 230-0045, Japan (E-mail: genome-res@gsc.riken.go.jp, URL: http://genome.gsc.riken.go.jp/, Tel: 81-45-503-9222, Fax: 81-45-503-9216)
Please visit our web site (http://genome.gsc.riken.go.jp/) for further details.
CDNA library was prepared and sequenced in Mouse Genome Encyclopedia Project of Genome Exploration Research Group in Riken Genomic Sciences Center and Genome Science Laboratory in RIKEN. Division of Experimental Animal Research in Riken contributed to prepare mouse tissues. First strand cDNA was primed with a primer

GenCore version 5.1.6
Copyright (c) 1993 - 2004 Compugen Ltd.

OM nucleic - nucleic search, using sw model

Run on: February 19, 2004, 04:41:17 ; Search time 350,095 Seconds
(without alignments)
9784.725 Million cell updates/sec

Title: US-10-042-431-46

Perfect score: 1269

Sequence: 1 atgttgaaacccctgtcaag.....gacgggtgtggcgcgtattg 1269

Scoring table: IDENTITY NUC

Gapop 10.0 , Gapext 1.0

Searched: 2552756 seqs, 1349719017 residues

Total number of hits satisfying chosen parameters: 5105512

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : N_Geneseq 19Jun03.*

- 1: /SIDSL1/gcgdata/geneseq/geneseq-emb1/NA1980.DAT.*
- 2: /SIDSL1/gcgdata/geneseq/geneseq-emb1/NA1981.DAT.*
- 3: /SIDSL1/gcgdata/geneseq/geneseq-emb1/NA1982.DAT.*
- 4: /SIDSL1/gcgdata/geneseq/geneseq-emb1/NA1983.DAT.*
- 5: /SIDSL1/gcgdata/geneseq/geneseq-emb1/NA1984.DAT.*
- 6: /SIDSL1/gcgdata/geneseq/geneseq-emb1/NA1985.DAT.*
- 7: /SIDSL1/gcgdata/geneseq/geneseq-emb1/NA1986.DAT.*
- 8: /SIDSL1/gcgdata/geneseq/geneseq-emb1/NA1987.DAT.*
- 9: /SIDSL1/gcgdata/geneseq/geneseq-emb1/NA1988.DAT.*
- 10: /SIDSL1/gcgdata/geneseq/geneseq-emb1/NA1989.DAT.*
- 11: /SIDSL1/gcgdata/geneseq/geneseq-emb1/NA1990.DAT.*
- 12: /SIDSL1/gcgdata/geneseq/geneseq-emb1/NA1991.DAT.*
- 13: /SIDSL1/gcgdata/geneseq/geneseq-emb1/NA1992.DAT.*
- 14: /SIDSL1/gcgdata/geneseq/geneseq-emb1/NA1993.DAT.*
- 15: /SIDSL1/gcgdata/geneseq/geneseq-emb1/NA1994.DAT.*
- 16: /SIDSL1/gcgdata/geneseq/geneseq-emb1/NA1995.DAT.*
- 17: /SIDSL1/gcgdata/geneseq/geneseq-emb1/NA1996.DAT.*
- 18: /SIDSL1/gcgdata/geneseq/geneseq-emb1/NA1997.DAT.*
- 19: /SIDSL1/gcgdata/geneseq/geneseq-emb1/NA1998.DAT.*
- 20: /SIDSL1/gcgdata/geneseq/geneseq-emb1/NA1999.DAT.*
- 21: /SIDSL1/gcgdata/geneseq/geneseq-emb1/NA2000.DAT.*
- 22: /SIDSL1/gcgdata/geneseq/geneseq-emb1/NA2001A.DAT.*
- 23: /SIDSL1/gcgdata/geneseq/geneseq-emb1/NA2001B.DAT.*
- 24: /SIDSL1/gcgdata/geneseq/geneseq-emb1/NA2002.DAT.*
- 25: /SIDSL1/gcgdata/geneseq/geneseq-emb1/NA2003.DAT.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	1269	100.0	1269	22	AAF45132 Human TANGO 294 OR
2	1269	100.0	2044	22	AAF45131 Human TANGO 294 CD
3	1269	100.0	2044	25	ABX94106 cDNA encoding huma
4	1267.4	99.9	2035	22	AAI19226 Human CG162 (or C5
5	1235.4	97.4	1384	22	AAI19220 Human CG162 (or C5
6	1122	88.4	1206	24	ABX55356 cDNA encoding huma
7	1122	88.4	1210	24	ABK85659 Human cDNA encodin
8	1104.8	87.1	1272	24	ABK90871 cDNA encoding huma

9	752	59.3	1116	24	ABK86570 Human cDNA encodin
10	527.8	41.6	2481	24	ABN97383 Gene #3881 used to
11	526.2	41.5	2493	22	AAH02906 Human shear stress
12	489.2	38.6	1378	14	AAQ02310 RGL precursor. Or
13	470	37.0	1367	7	AAH06085 Sequence encoding
14	470	37.0	1367	7	AAH06085 Sequence encoding
15	470	37.0	1367	17	AAH58916 Human stomach cell
16	468.4	36.9	1695	22	AAH57420 Human gastric lipa
17	464.4	36.6	1336	6	AAH50385 Rat prelingual lip
18	456.8	36.0	1411	24	ABL57144 Kid goat pregastr
19	454.2	35.8	1194	24	ABL27803 Human lysosomal ac
20	454	35.8	1228	24	ABK12385 cDNA encoding huma
21	450.8	35.5	1284	24	ABK85978 DNA encoding huma
22	449.8	35.4	8324	24	ABL57145 yeast YE-1 express
23	448.4	35.3	1473	24	ABK12388 cDNA encoding huma
24	446.8	35.2	1185	24	ABK12388 Human triacylglyce
25	444.2	35.0	1528	15	AAQ68388 Canine gastric lip
26	444.2	35.0	1528	17	AAH58915 Dog gastric lipase
27	444.2	35.0	1531	15	AAQ68389 Canine gastric lip
28	443.6	35.0	1200	22	AAH78205 Nucleotide sequenc
29	424.8	33.5	1260	24	ABK12386 DNA encoding huma
30	419.8	33.1	1086	25	ACC44783 Human triacylglyce
31	414.2	32.6	1360	24	ABL47480 Human lysosomal ac
32	414.2	32.6	1360	25	ABX14880 Human lipase cDNA.
33	374.4	29.5	996	25	ACC48563 Human dithp secret
34	346.4	27.3	1138	24	ABQ82326 Human NOV1 encodin
35	332.8	26.2	1098	24	AAH27802 Human lysosomal ac
36	319.4	25.2	1143	24	ABK12386 cDNA encoding huma
37	315.4	24.9	828	24	ABK12386 Human triacylglyce
38	300.4	23.7	1377	24	ABK12386 DNA encoding huma
39	280	22.9	842	24	AAH57102 Human triacylglyce
40	286	22.5	699	22	AAF28679 Human protein HP03
41	286	22.5	1308	22	AAF28689 Human protein HP03
42	282.8	22.3	802	22	AAI97406 Human neuroblastom
43	267.4	21.1	18554	24	ABX55357 Genomic DNA encodi
44	257.8	20.3	617	21	AAA44349 Human secreted exp
45	244.6	19.3	1971	23	AAH591026 DNA encoding novel

ALIGNMENTS

RESULT 1
ID AAF45132 standard; cDNA; 1269 BP.
XX AAF45132;
XX AAF45132;
DT 30-MAR-2001 (first entry)
XX Human TANGO 294 ORF.

XX Gene therapy; TANGO protein; INTERCEPT protein; neurological disorder;
XX central nervous system; focal brain disorder; bipolar affective disorder;
XX global-diffuse cerebral disorder; cerebrovascular; Alzheimer's disease;
XX senile dementia; Huntington's disease; amyotrophic lateral sclerosis;
XX Parkinson's; Gilles de la Tourette's syndrome; hypertension; sleep;
XX neuropsychiatric; psychoactive substance use; anxiety; ss.
OS Homo sapiens.
XX
XX WO200077239-A2.
XX
XX 21-DEC-2000.
XX
XX 24-MAY-2000; 2000WO-US14858.
XX
XX 14-JUN-1999; 99US-0333159.
XX
XX (MILL-) MILLENNIUM PHARM INC.
XX
XX PI McCarthy SA, Fraser CC, Sharp JD, Barnes TM;

DR WPI: 2001-032313/04.
 DR P-PSDB; AAB66065.
 XX
 PT TANGO and INTERCEPT nucleic acids, proteins, and antibodies, useful for
 PT screening assays and diagnostic assays and for the treatment of
 PT neurological diseases such as Alzheimer's, Parkinson's and Huntington's
 PT disease -
 XX
 XX Claim 1; Fig 6; 359pp; English.
 XX
 CC The present invention relates to TANGO or INTERCEPT proteins and coding
 CC sequences (see AAF45131-F45136 and AAF45138-F45139 and AAB66031-B66057,
 CC AAB66064-B66083 and AAB66085). The TANGO/INTERCEPT proteins and coding
 CC sequences are useful for the treatment of neurological disorders such as
 CC central nervous system (CNS) disorders, CNS-related disorders, focal
 CC brain disorders, global-diffuse cerebral disorders and other
 CC neurological and cerebrovascular disorders. The CNS disorders include
 CC Alzheimer's disease, senile dementia, Huntington's disease, amyotrophic
 CC lateral sclerosis, Parkinson's, Gilles de la Tourette's syndrome,
 CC autonomic function disorders such as hypertension and sleep disorders,
 CC neuropsychiatric disorders, psychoactive substance use disorders,
 CC anxiety, and bipolar affective disorder.
 XX
 SQ Sequence 1269 BP; 358 A; 268 C; 309 G; 334 T; 0 other;
 Query Match 100.0%; Score 1269; DB 22; Length 1269;
 Best Local Similarity 100.0%; Pred. No. 0;
 Matches 1269; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 1 ATGTTGGAAACCTTGTCAGACAGTGGATGTCACACAGATGGAAATGGCTTCG 60
 1 ATGTTGGAAACCTTGTCAGACAGTGGATGTCACACAGATGGAAATGGCTTCG 60
 61 ATTCTGGTGGCGTATATGTTCCAGAGAAATGTAATTCAGTACATATGCAACTAAAGCT 120
 61 ATTCTGGTGGCGTATATGTTCCAGAGAAATGTAATTCAGTACATATGCAACTAAAGCT 120
 121 GTGACCCGACAGCATTCATGAATATATAGTGAATCATTCACATCAAGGCTATCCCTGT 180
 121 GTGACCCGACAGCATTCATGAATATATAGTGAATCATTCACATCAAGGCTATCCCTGT 180
 181 GAGGAATATCAAGTCGCACTGAGATGGATGATATCTTCTGTTAAACAGGATTCCTCGA 240
 181 GAGGAATATCAAGTCGCACTGAGATGGATGATATCTTCTGTTAAACAGGATTCCTCGA 240
 241 GGCTAGTGCACCTCAAGAGACAGGTTCCAGGCTGTGGTGTACTGACAGATGGCCTA 300
 241 GGCTAGTGCACCTCAAGAGACAGGTTCCAGGCTGTGGTGTACTGACAGATGGCCTA 300
 301 GTTGGAGGTGCTAGCACTGGATTTCCAACTGCCCAACATAGCTGGGCTTCATCTG 360
 301 GTTGGAGGTGCTAGCACTGGATTTCCAACTGCCCAACATAGCTGGGCTTCATCTG 360
 361 GCAGATGCTGGTGTGAGTGGGAAACAGCAGGAGGAAACGCTGGTCTCGAAA 420
 361 GCAGATGCTGGTGTGAGTGGGAAACAGCAGGAGGAAACGCTGGTCTCGAAA 420
 421 CACAGACACTCTCCATAGACAGATGATGTTCTGGGCTTCATGATGATGATGGCT 480
 421 CACAGACACTCTCCATAGACAGATGATGTTCTGGGCTTCATGATGATGATGGCT 480
 481 AGGTTTGACCTTCTGCACTGATATACTTTATTTGCAAGAAACGGGCCAGGAAAGATC 540
 481 AGGTTTGACCTTCTGCACTGATATACTTTATTTGCAAGAAACGGGCCAGGAAAGATC 540
 541 TATTATGTGGCTATTACAGGGGACCAACATGGGCTTTATGATTTTCCACCATGCCA 600
 541 TATTATGTGGCTATTACAGGGGACCAACATGGGCTTTATGATTTTCCACCATGCCA 600
 601 GAGCTGGCTCAGAAAATCAAAATGTAATTTGCTTTAGCACCCATAGCCATGTTAAGCAT 660
 601 GAGCTGGCTCAGAAAATCAAAATGTAATTTGCTTTAGCACCCATAGCCATGTTAAGCAT 660

QY 661 GC AAAAGCCCGGACCAAAATTTTGTGTCGCAGATATGATGATCAAGGATTTGTT 720
 DB 661 GC AAAAGCCCGGACCAAAATTTTGTGTCGCAGATATGATGATCAAGGATTTGTT 720
 QY 721 GGCAAAAAGAAATTTCTGTATCAGACCAAGTTCTCAGCAACTTGTATTACCTTTG 780
 DB 721 GGCAAAAAGAAATTTCTGTATCAGACCAAGTTCTCAGCAACTTGTATTACCTTTG 780
 QY 781 GSCCAGGTGATTTCTGATCAGATTTCTAGTATATATATCTCTGCTGGTGGATTCAAC 840
 DB 781 GSCCAGGTGATTTCTGATCAGATTTCTAGTATATATATCTCTGCTGGTGGATTCAAC 840
 QY 841 ACCAACAAATATGAACATGAGCCGAGCAAGTGTATATGCTGCCACACTCTTGTGGAACA 900
 DB 841 ACCAACAAATATGAACATGAGCCGAGCAAGTGTATATGCTGCCACACTCTTGTGGAACA 900
 QY 901 TCTGTGCAAAATATTTCTACACTGGAGCCGAGGAGTGAATTTCTGGTGAATTCGGGCTTT 960
 DB 901 TCTGTGCAAAATATTTCTACACTGGAGCCGAGGAGTGAATTTCTGGTGAATTCGGGCTTT 960
 QY 961 GACTGGGGGAGTGAGACCAAAATCTGGAATAATGCAATCAGCCAACTCTCTGTAAGGTAC 1020
 DB 961 GACTGGGGGAGTGAGACCAAAATCTGGAATAATGCAATCAGCCAACTCTCTGTAAGGTAC 1020
 QY 1021 AGAGTCAGATATGAGCGTCCCTACAGCAATGTGACAGAGGTCAGGACTGGCTTTCA 1080
 DB 1021 AGAGTCAGATATGAGCGTCCCTACAGCAATGTGACAGAGGTCAGGACTGGCTTTCA 1080
 QY 1081 AATCCAGAAGACGTGAAAATGCTGCTCTGAGGTGACCAACCTCATCTACCAATAAGAAAT 1140
 DB 1081 AATCCAGAAGACGTGAAAATGCTGCTCTGAGGTGACCAACCTCATCTACCAATAAGAAAT 1140
 QY 1141 ATTCTGAAATGGGCTCAGTGGATTTCACTGCGGTTTGGATGCTCTCTACCGTATGTAC 1200
 DB 1141 ATTCTGAAATGGGCTCAGTGGATTTCACTGCGGTTTGGATGCTCTCTACCGTATGTAC 1200
 QY 1201 AATCAATATCATCTCATCTGATGACAGGAGGAGCAACCTTTCCAGGAGCGTGTGAG 1260
 DB 1201 AATCAATATCATCTCATCTGATGACAGGAGGAGCAACCTTTCCAGGAGCGTGTGAG 1260
 QY 1261 GCCGTATTG 1269
 DB 1261 GCCGTATTG 1269
 RESULT 2
 AAF45131
 ID AAF45131 standard; cDNA; 2044 BP.
 XX
 AC AAF45131;
 XX
 DT 30-MAR-2001 (first entry)
 XX
 DE Human TANGO 294 cDNA.
 XX
 KW Gene therapy; TANGO protein; INTERCEPT protein; neurological disorder;
 KW central nervous system; focal brain disorder; bipolar affective disorder;
 KW global-diffuse cerebral disorder; cerebrovascular; Alzheimer's disease;
 KW senile dementia; Huntington's disease; amyotrophic lateral sclerosis;
 KW Parkinson's; Gilles de la Tourette's syndrome; hypertension; sleep;
 KW neuropsychiatric; psychoactive substance use; anxiety; ss.
 OS Homo sapiens.
 XX
 XX WO200077239-A2.
 XX
 PD 21-DEC-2000.
 XX
 PF 24-MAY-2000; 2000WO-US14858.
 XX
 PR 14-JUN-1999; 99US-0333159.
 XX
 PA (MILL-) MILLENNIUM PHARM INC.

CDS 126..1397
 /cag= a
 /product= "TANGO 294"
 /note= "The ORF given as SEQ ID No:46 is specifically
 claimed in Claim 2"
 US2002182675-A1.
 05-DEC-2002.
 25-OCT-2001; 2001US-0042431.
 14-JUN-1999; 99US-0333159.
 24-MAY-2000; 2000US-0578063.
 (MILL-) MILLENNIUM PHARM INC.
 PI McCarthy SA, Fraser CC, Sharp JD, Barnes TM;
 WPI; 2003-328617/31.
 DR P-PSDB; ABU08369.
 XX
 PT New TANGO and INTERCEPT proteins, useful as modulating agents in
 PT regulating a variety of cellular processes, in chromosome mapping, in
 PT tissue typing, and in forensic biology -
 XX
 PS Claim 2; Fig 6A-6C; 232pp; English.
 CC The present invention relates to the isolation of novel TANGO or
 CC INTERCEPT proteins, and the polynucleotide sequences encoding them.
 CC The invention discloses sequences for human TANGO 202, TANGO 234,
 CC TANGO 265, TANGO 273, TANGO 286, TANGO 294 and INTERCEPT 296. The
 CC sequences for murine TANGO 202 and TANGO 273 are also provided. The
 CC TANGO polypeptides and the polynucleotide sequences encoding them are
 CC useful as modulating agents in regulating a variety of cellular
 CC processes. The polynucleotide sequences are useful as primers or
 CC hybridisation probes for the detection of nucleic acids encoding
 CC TANGO polypeptides, chromosome mapping, tissue typing, and in
 CC forensic biology. TANGO 202 can be used to diagnose, prevent or
 CC treat disorders relating to aberrant cellular protease activity,
 CC inappropriate interaction of cells with mediators, inappropriate
 CC development, and blood and haematopoietic cell-related disorders.
 CC TANGO 234 can be used to modulate growth, proliferation, survival,
 CC differentiation, and activity of gamma delta T-cells. TANGO 265 can
 CC be used to prevent, diagnose and treat disorders characterised by
 CC aberrant organisation or development of a tissue or organ, and for
 CC modulating differentiation of cells of the immune system. TANGO 273
 CC is useful for diagnosing, treating or preventing e.g. metabolic,
 CC homeostatic and developmental bone disorders (e.g. osteoporosis),
 CC and bacterial infection. TANGO 286 is useful for treating or
 CC preventing e.g. anaemia, thrombocytopaenia, renal failure or
 CC liver disease. TANGO 294 is useful for treating or preventing
 CC e.g. cystic fibrosis or obesity. INTERCEPT 296 is useful for
 CC diagnosing, treating or preventing e.g. cancers, bronchitis, cystic
 CC fibrosis, asthma, emphysema, pulmonary oedema, or adult and infant
 CC respiratory distress syndromes. The present sequence encodes
 CC human TANGO 294.
 XX
 SQ Sequence 2044 BP; 618 A; 401 C; 460 G; 565 T; 0 other;
 Query Match 100.0%; Score 1269; DB 25; Length 2044;
 Best Local Similarity 100.0%; Pred. No. 0;
 Matches 1269; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 QY 1 ATGTTGGAACTTGTCCAGACAGTGTCTCCACAGAGATGGAAATGGCTTCTG 60
 DB 126 ATGTTGGAACTTGTCCAGACAGTGTCTCCACAGAGATGGAAATGGCTTCTG 185
 QY 61 ATTCTGGGGGTATATGTTCCAGAGAAATGTGAATTCAGTACATATGCCAACTAAAGCT 120
 DB 186 ATTCTGGGGGTATATGTTCCAGAGAAATGTGAATTCAGTACATATGCCAACTAAAGCT 245
 QY 121 GTGAGCCCAAGCAATTCATGATATAGTGAATATCCAAATCAATCAAGGCTATCCCTGT 180

246 GTGAGCCCAAGCAATTCATGATATAGTGAATATCCAAATCAATCAAGGCTATCCCTGT 305
 181 GAGGAATATGAGTCCGCACTGAAGATGGTATATCTTCTCTGTTTAAAGGATTCCTCGA 240
 306 GAGGAATATGAGTCCGCACTGAAGATGGTATATCTTCTCTGTTTAAAGGATTCCTCGA 365
 241 GGCCTAGTGCACCTAAGAAAGACAGGTTCCAGGCTCTGTGTGTTTACTGACAGCATGGCCTA 300
 366 GGCCTAGTGCACCTAAGAAAGACAGGTTCCAGGCTCTGTGTGTTTACTGACAGCATGGCCTA 425
 301 GTTGGAGGTGCTAGCAATTCGATTTCCACCTGCGCCACAAATAGCTTGGCTTCATCTG 360
 426 GTTGGAGGTGCTAGCAATTCGATTTCCACCTGCGCCACAAATAGCTTGGCTTCATCTG 485
 361 GCAGATGCTGTTTGTGACGTGTGATGGGAAACAGAGGGGAAACGCTGCTTCGAAAA 420
 486 GCAGATGCTGTTTGTGACGTGTGATGGGAAACAGAGGGGAAACGCTGCTTCGAAAA 545
 421 CACAAGACATCTCCATAGACCAAGATGAGTTCTGGGCTTTCAGTTATGATGATGATGGCT 480
 546 CACAAGACATCTCCATAGACCAAGATGAGTTCTGGGCTTTCAGTTATGATGATGATGGCT 605
 481 AGTTTTCAGCTTCTGTCAGTGTATTAATTTTTCAGAAAAACGCGCCAGGAAAGATC 540
 606 AGTTTTCAGCTTCTGTCAGTGTATTAATTTTTCAGAAAAACGCGCCAGGAAAGATC 665
 541 TATTATGTCGGCTATTTCAGAGGCAACCATGGGCTTATTCATGATTTTCCACATGCCA 600
 666 TATTATGTCGGCTATTTCAGAGGCAACCATGGGCTTATTCATGATTTTCCACATGCCA 725
 601 GAGTGGCTCAGAAATCAAAATGATTTTCTTACGCCCATAGCCATAGCCATGTTAGCAT 660
 726 GAGTGGCTCAGAAATCAAAATGATTTTCTTACGCCCATAGCCATAGCCATGTTAGCAT 785
 661 GCAAAAAAGCCCGGAGCAAAATTTTCTGTCGCCAGATATGATGATCAAGGATTTGTT 720
 786 GCAAAAAAGCCCGGAGCAAAATTTTCTGTCGCCAGATATGATGATCAAGGATTTGTT 845
 721 GCGAAAAAGAAATTTCTGATCAGACAGATTTCTCAGACACTTGTATTACCTTTGT 780
 846 GCGAAAAAGAAATTTCTGATCAGACAGATTTCTCAGACACTTGTATTACCTTTGT 905
 781 GCGAGGTGATTTCTGATCAGATTTGTAGTAATATCATGTTTCTTGGGTGGATTCAC 840
 906 GCGAGGTGATTTCTGATCAGATTTGTAGTAATATCATGTTTCTTGGGTGGATTCAC 965
 841 ACCAACATATGAACATGAGCCGAGCAAGTGTATATGCTGCCACACTCTTGTGGAACA 900
 966 ACCAACATATGAACATGAGCCGAGCAAGTGTATATGCTGCCACACTCTTGTGGAACA 1025
 901 TCTGTGCAAAATATTTCTACCTGAGCCAGCAGTGAATTTCTGTGAATCTCGGGCATTT 960
 1026 TCTGTGCAAAATATTTCTACCTGAGCCAGCAGTGAATTTCTGTGAATCTCGGGCATTT 1085
 961 GACTGGGGGAGTGAGACCAAAATCTGGAATAATGCAATCAGCCAACTCTCTGTAAGTAC 1020
 1086 GACTGGGGGAGTGAGACCAAAATCTGGAATAATGCAATCAGCCAACTCTCTGTAAGTAC 1145
 1021 AGAGTCAGAGATATGACGGTCCCTACAGCATGTGGACAGAGTCAAGGCTGAGGCTTCA 1080
 1146 AGAGTCAGAGATATGACGGTCCCTACAGCATGTGGACAGAGTCAAGGCTGAGGCTTCA 1205
 1081 AATCCAGAGAGCTGGAATGCTGCTCTCTGAGGTGACCACTCATCTACCATAGAAT 1140
 1206 AATCCAGAGAGCTGGAATGCTGCTCTCTGAGGTGACCACTCATCTACCATAGAAT 1265
 1141 ATTCTGTAATGGGCTCAGTGGATTTCTATCTGGGGTTTGGATGCTCTCCACCGTATGTAC 1200
 1266 ATTCTGTAATGGGCTCAGTGGATTTCTATCTGGGGTTTGGATGCTCTCCACCGTATGTAC 1325
 1201 AATGAATATCCATCTGATCAGGAGGAGGAGCAACCTTTCCAGGAGCGGTGTGAG 1260

Db 1326 ATGGAATCATCTCTGATGACGAGGAGGACCAACCTTCCAGGAGCGGTGTGAG 1385

Qy 1261 GCGGTATTG 1269
 Db 1386 GCGGTATTG 1394

RESULT 4

AAID19226
 ID AAD19226 standard; DNA; 2035 BP.

XX AC AAD19226;

DT 18-DEC-2001 (first entry)

XX DE Human CG162 (or C59) lipase DNA #2.

XX KW Human; apolipoprotein; lipase; lipoprotein receptor; ALLr; angina;
 KW cardiovascular disease; lipid metabolism; myocardial infarction;
 KW cerebral ischaemia; arterial thrombosis; thrombolytic; antilipemic;
 KW coronary artery thrombosis; cerebral artery thrombosis; stroke;
 KW intracardiac thrombosis; gene therapy; cardiovascular; vasodilator;
 KW neuroprotectant; cerebroprotective; db.

XX OS Homo sapiens.

XX FH Key Location/Qualifiers
 XX CDS 189..1418

FT /tag= a
 FT /product= "Human CG162 (or C59) lipase protein"
 FT 189..245

FT sig_peptide

FT /tag= b

FT mat_peptide

FT 246..1415

FT /tag= c

FT /product= "Human mature CG162 (or C59) lipase protein"

XX WO200179446-A2.

XX PD 25-OCT-2001.

XX 16-APR-2001; 2001WO-US12529.

XX 14-APR-2000; 2000US-197137P.

XX 20-JUN-2000; 2000US-0598042.

XX 03-AUG-2000; 2000US-0631451.

XX 22-SEP-2000; 2000US-0667298.

XX 17-NOV-2000; 2000US-0714936.

XX (HYSE-) HYSEQ INC.

XX PI Ballinger DG, Loeb D, Montgomery JR, Tang TY, Zhou P, Goodrich R;
 PI Liu C, Asundi V, Zhao QA, Wehrman T, Drmanac RT, Ren F, Qian XB;
 PI Wang D;

XX WPI: 2001-611724/70.

XX P-PSDB; AAE11931.

XX Nucleic acids encoding human apolipoproteins, lipases, and lipoprotein
 PT receptor polypeptides, useful for preventing diagnosing and treating
 PT lipid metabolism disorders, thrombosis and cardiovascular diseases -

XX Claim 1; Page 183-185; 266pp; English.

XX The invention relates to polynucleotides encoding proteins CG122, CG179,
 CC CG95, CG121, CG162, CG27, CG153 and CG168 which are related to proteins
 CC involved in lipid metabolism and cardiovascular disease such as human
 CC apolipoproteins, lipases and lipoprotein receptor proteins. These DNA
 CC and protein sequences are useful for treating or preventing disorders
 CC associated with apolipoproteins, lipases and lipoprotein receptor (ALLr)
 CC expression and for treating lipid metabolism, cardiovascular diseases
 CC and thrombosis. Antibodies against these proteins are useful for
 CC determining the presence of or predisposition to a disease associated

CC with altered levels of these sequences. ALLr polypeptides are also
 CC useful for identifying agents (agonists and antagonists) that bind to
 CC them and cells expressing ALLr proteins are useful for identifying a
 CC therapeutic agent for use in treatment of a pathology related to
 CC aberrant expression or physiological interactions of this polypeptide.
 CC Vectors comprising these DNA and protein sequences are also useful for
 CC producing ALLr proteins. The nucleic acids and polypeptides of the
 CC invention are also useful for the treatment of occlusive cardiovascular
 CC diseases, myocardial infarction, cerebral ischaemia, angina, arterial
 CC thrombosis, coronary artery thrombosis and cerebral artery thrombosis
 CC or intracardiac thrombosis and stroke. The nucleotides of the invention
 CC are used in gene therapy. The present sequence is human CG162 (or C59)
 CC lipase DNA.

XX Sequence 2035 BP; 508 A; 395 C; 456 G; 576 T; 0 other;

Query Match 99.9%; Score 1267.4; DB 22; Length 2035;
 Best Local Similarity 99.9%; Pred. No. 0;
 Matches 1268; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

Qy 1 ATGTTGGAAACCTTGTCAAGACAGTGGATTGTCTCACACAGATGCAATGGCTTCG 60

Db 147 ATGTTGGAAACCTTGTCAAGACAGTGGATTGTCTCACACAGATGCAATGGCTTCG 206

Qy 61 ATTTCTGGTGGCGTATATGTTCCAGAGAAATGGAATTCAGTACATATGCCAACTAAAGCT 120

Db 207 ATTTCTGGTGGCGTATATGTTCCAGAGAAATGGAATTCAGTACATATGCCAACTAAAGCT 266

Qy 121 GTGACCCAGACCATTCATGATATAGTGAATCCTCCACATCCACATCAGGCTATCCCTGT 180

Db 267 GTGACCCAGACCATTCATGATATAGTGAATCCTCCACATCCACATCAGGCTATCCCTGT 326

Qy 181 GAGGAATATGAGTCCGAACTGAAATGGAATTCAGTACATATGCCAACTAAAGCT 240

Db 327 GAGGAATATGAGTCCGAACTGAAATGGAATTCAGTACATATGCCAACTAAAGCT 386

Qy 241 GGCTTAGTGCAACCTAAGAAAGACAGGTTCCAGGCTGTGGTGTACTGAGCATGGCCTTA 300

Db 387 GGCTTAGTGCAACCTAAGAAAGACAGGTTCCAGGCTGTGGTGTACTGAGCATGGCCTTA 446

Qy 301 GTTGGAGGTGCTAGCAACTGGATTTCACACCTGCCCAACAATAGCTGGGCTTCATTCTG 360

Db 447 GTTGGAGGTGCTAGCAACTGGATTTCACACCTGCCCAACAATAGCTGGGCTTCATTCTG 506

Qy 361 GCAGATGCTGTTTTCACGTGTGGATGGGAAACAGCAGGGGAAACGCTGTGTCGAAA 420

Db 507 GCAGATGCTGTTTTCACGTGTGGATGGGAAACAGCAGGGGAAACGCTGTGTCGAAA 566

Qy 421 CACAAGACACTCTCCATAGACCAAGATGATCTGGGCTTCAGTTATGATGAGATGGCT 480

Db 567 CACAAGACACTCTCCATAGACCAAGATGATCTGGGCTTCAGTTATGATGAGATGGCT 626

Qy 481 AGTTTGGACCTTCCTGCGATGATAAACTTTATTTTGACAGAAACGGGGCAGGAAAGATC 540

Db 627 AGTTTGGACCTTCCTGCGATGATAAACTTTATTTTGACAGAAACGGGGCAGGAAAGATC 686

Qy 541 TATTATGTGGCTATTTCACAGGCAACCATGGGCTTTATTCATTTTCCACCATGCCA 600

Db 687 TATTATGTGGCTATTTCACAGGCAACCATGGGCTTTATTCATTTTCCACCATGCCA 746

Qy 601 GAGCTGGCTCAGAAAATCAAAATGATTTTGTAGACCCCATAGCCACTGTAAAGCAT 660

Db 747 GAGCTGGCTCAGAAAATCAAAATGATTTTGTAGACCCCATAGCCACTGTAAAGCAT 806

Qy 661 GCAAAAGCCCGGGACCAAAATTTTGTGTCGACAGATATGATGATCAAGGATTTGTTT 720

Db 807 GCAAAAGCCCGGGACCAAAATTTTGTGTCGACAGATATGATGATCAAGGATTTGTTT 866

Qy 721 GGCAGAAAAGAAATTTCTGTATCAGACCAATTTCTCAGACAACTTCTTATTTACCTTCT 780

Db 867 GGCAGAAAAGAAATTTCTGTATCAGACCAATTTCTCAGACAACTTCTTATTTACCTTCT 926

Qy 781 GGCAGGATGATTTCTTGATCAGATTTTGTAGTAATATCATCTTCTTCTGGTGGATTCAC 840

QY 541 TATTATGCGCTATTCACAGGCGCACCCATCGGCTTTATTGCAATTTCCACCATGCCA 600
 Db 587 TATTATGCGGCTATTCACAGGCGCACCCATCGGCTTTATTGCAATTTCCACCATGCCA 746
 QY 601 GAGCTGGCTCAGAAATCAAAATGTAATTTGCTTTAGCACCATAGCCACTGTTAAGCAT 660
 Db 747 GAGCTGGCTCAGAAATCAAAATGTAATTTGCTTTAGCACCATAGCCACTGTTAAGCAT 806
 QY 561 GCAAAAGCCCGGCGCAAAATTTGCTGCTCCAGATATGATCATCAAGGATGTTT 720
 Db 807 GCAAAAGCCCGGCGCAAAATTTGCTGCTCCAGATATGATCATCAAGGATGTTT 866
 QY 721 GGCACAAAAGAAATTTCTGTATCAGACAGATTTCTCAGACAACTTGTATTACCTTTGT 780
 Db 867 GGCACAAAAGAAATTTCTGTATCAGACAGATTTCTCAGACAACTTGTATTACCTTTGT 926
 QY 781 GGCAGAGTGATCTTGATCAGATTTGTAATATCATCATGATTTCTGCGGTGATTCAC 840
 Db 927 GGCAGAGTGATTTGTATCAGATTTGTAATATCATCATGATTTCTGCGGTGATTCAC 986
 QY 841 ACCAAATATGAATGAGCCGAGCAAGTGTATATGCTGCCACACACTCTTGTGGAACA 900
 Db 987 ACCAAATATGAATGAGCCGAGCAAGTGTATATGCTGCCACACACTCTTGTGGAACA 1046
 QY 901 TCTGTGCAAAATATTTACTGAGCGAGGAGTGAATTTCTGGTGAATCTCCGGCATTT 960
 Db 1047 TCTGTGCAAAATATTTACTGAGCGAGGAGTGAATTTCTGGTGAATCTCCGGCATTT 1106
 QY 961 GACTGGGGAGTGAGACCAAAATCTGAAATATGCAATCAGCCAACTCTCTGTAAGTAC 1020
 Db 1107 GACTGGGGAGTGAGACCAAAATCTGAAATATGCAATCAGCCAACTCTCTGTAAGTAC 1166
 QY 1021 AGAGTCAGAGATATGACGGTCCCTACAGCAATGTGAGAGAGTCAAGCTGGCTTTCA 1080
 Db 1167 AGAGTCAGAGATATGACGGTCCCTACAGCAATGTGAGAGAGTCAAGCTGGCTTTCA 1226
 QY 1081 AATCCAGAGACGTGAAATGCTGCTCTCAGAGTGACCACTCATCTACCAATGAAT 1140
 Db 1227 AATCCAGAGACGTGAAATGCTGCTCTCAGAGTGACCACTCATCTACCAATGAAT 1286
 QY 1141 ATTCTGAATGGGCTCAGTGGATTTCACTGGGGTTTGGATGCTCTCTCAACGATGATAC 1200
 Db 1287 ATTCTGAATGGGCTCAGTGGATTTCACTGGGGTTTGGATGCTCTCTCAACGATGATAC 1346
 QY 1201 AATGAATATCCATCTGATGAGGAGGAGGACCA 1237
 Db 1347 AATGAATATCCATCTGATGAGGAGGAGGACCA 1383

RESULT 6

ID ABS55356 standard; cDNA; 1206 BP.
 XX AC ABS55356;
 XX AC ABS55356;
 DT 07-JAN-2003 (first entry)
 DE cDNA encoding human lipase protein.
 XX Human; lipase; cancer; colon; kidney; skin; brain; testis; chromosome 10;
 XX Burkitt's lymphoma; cytostatic; gene therapy; lipase-inhibitor; gene; ss.
 OS Homo sapiens.
 PH Key Location/Qualifiers
 FT 1..3 /tag= a
 FT 5'UTR
 FT CDS
 FT 4..1200 /tag= b
 FT /product= "Human lipase protein"
 FT 1201..1206
 FT 3'UTR
 FT /tag= c
 XX

PN WO200274975-A2.
 XX 26-SEP-2002.
 XX 18-MAR-2002; 2002WO-US08035.
 XX 20-MAR-2001; 2001US-0811825.
 XX (PEKE) PE CORP NY.
 PA Yan C, Di Francesco V, Beasley EM;
 PI WPI; 2002-750560/81.
 DR P-PSDB; ABG71000.
 XX New isolated human lipase peptides and encoding nucleic acids, useful
 PT for diagnosing and treating disorders mediated by human lipase
 PT proteins, e.g. cancer of the colon, kidneys, skin and testis, and
 PT Burkitt's lymphoma -
 XX Claim 4; Fig 1; 79pp; English.
 XX The present invention relates to a new lipase peptide. The methods and
 CC compositions of the present invention are useful for diagnosing and
 CC treating disorders mediated by the human lipase protein, such as cancer
 CC of the colon, kidneys, skin, brain and testis, and Burkitt's lymphoma.
 CC The present nucleic acid sequence represents the human lipase gene
 CC located on chromosome 10. This sequence encodes the human lipase protein
 CC of the invention.
 XX Sequence 1206 BP; 339 A; 260 C; 295 G; 312 T; 0 other;
 PS Query Match 88.4%; Score 1122; DB 24; Length 1206;
 XX Best Local Similarity 100.0%; Pred. No. 0;
 XX Matches 1122; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 QY 148 AGTGAATATCATCAACATCAAGGCTATCCCTGTGAGGAATATGAGTCGCAACTGAAGAT 207
 Db 76 AGTGAATATCATCAACATCAAGGCTATCCCTGTGAGGAATATGAGTCGCAACTGAAGAT 135
 QY 208 GGGTATATCTTTCTGTTAAAGGATTCCTCGAGGCTAGTGCACCTAAGAGACAGGT 267
 Db 136 GGGTATATCTTTCTGTTAAAGGATTCCTCGAGGCTAGTGCACCTAAGAGACAGGT 195
 QY 268 TCCAGGCTGTGCTGTACTGAGCATGCGCTAGTTGAGGCTGCTAGCACTGGATTTC 327
 Db 196 TCCAGGCTGTGCTGTACTGAGCATGCGCTAGTTGAGGCTGCTAGCACTGGATTTC 255
 QY 328 AACCTGCCCAACAATAGCTGGCTTCATTTCTGGCAGATGCTGTTTGAAGTGGATG 387
 Db 256 AACCTGCCCAACAATAGCTGGCTTCATTTCTGGCAGATGCTGTTTGAAGTGGATG 315
 QY 388 GGGACAGCAGGGGAAACGCTGCTCGAAACACAGACACTCTCCATAGACCAAGAT 447
 Db 316 GGGACAGCAGGGGAAACGCTGCTCGAAACACAGACACTCTCCATAGACCAAGAT 375
 QY 448 GAGTCTGGGCTTTTCAGTTATGATGAGATGGCTAGGTTTGAACCTTCTGCAAGTATAAC 507
 Db 376 GAGTCTGGGCTTTTCAGTTATGATGAGATGGCTAGGTTTGAACCTTCTGCAAGTATAAC 435
 QY 508 TTTATTTTGCAGAAACGGGCGCAGGAAGATCTATTATGCTGGCTATTTCACAGGAC 567
 Db 436 TTTATTTTGCAGAAACGGGCGCAGGAAGATCTATTATGCTGGCTATTTCACAGGAC 495
 QY 568 ACCATGGGCTTTTATGCTATTTTCCACCATGCGAGCTGGCTCAGAAATCAAAATGTAT 627
 Db 496 ACCATGGGCTTTTATGCTATTTTCCACCATGCGAGCTGGCTCAGAAATCAAAATGTAT 555
 QY 628 TTTGCTTTAGACCCCATAGCCACTGTTAAGCATGCAAAAAGCCCGGACCAATTTTGG 687
 Db 556 TTTGCTTTAGACCCCATAGCCACTGTTAAGCATGCAAAAAGCCCGGACCAATTTTGG 615
 QY 688 TTGCTGCCAGATATGATCAAGGATTTGTTTGGCAAAAAGAAATTTCTGTATCAGACC 747

235	CCTGAGCCCTAGTGCACCTTAAGAAGCAGGTTCCAGGCCCTGTGGTGTACTGAGCAT	294	QY
236		295	QY
237		296	QY
214	CCTCAGGCCCTAGTGCACCTTAAGAAGCAGGTTCCAGGCCCTGTGGTGTACTGAGCAT	273	Db
215		274	Db
216		275	Db
295	GGCCTAGTTGGAGGTCTAGCAACTGGATTTTCCAACTCTGCCCAAACAATAGCCTGGGCTTC	354	QY
296		355	QY
274	GGCCTAGTTGGAGGTCTAGCAACTGGATTTTCCAACTCTGCCCAAACAATAGCCTGGGCTTC	333	Db
275		334	Db
355	ATTCTGGCAGATGCTGGTTTTGAAGTGTGGATGGGGAACAGCAGGGGAAAAGCCTGGTCT	414	QY
356		415	QY
334	ATTCTGGCAGATGCTGGTTTTGAAGTGTGGATGGGGAACAGCAGGGGAAAAGCCTGGTCT	393	Db
335		394	Db
415	CGAAAAACAAGACACTCTCCCATAGACCAAGATGAGTTCTGGGCTTTTCAGTTTATGATGAG	474	QY
416		475	QY
394	CGAAAAACAAGACACTCTCCCATAGACCAAGATGAGTTCTGGGCTTTTCAGTTTATGATGAG	453	Db
395		454	Db
475	ATGGCTAGTTTGACCTTCTGCACTGATTAACACTTTATTTTCAGAAAACGGCCAGGAA	534	QY
476		535	QY
454	ATGGCTAGTTTGACCTTCTGCACTGATTAACACTTTATTTTCAGAAAACGGCCAGGAA	513	Db
455		514	Db
535	AAGACTATTATGTGCGCTATTACAGGGCCACCACTGGGCTTTATTCATTTTCCACC	594	QY
536		595	QY
514	AAGACTATTATGTGCGCTATTACAGGGCCACCACTGGGCTTTATTCATTTTCCACC	573	Db
515		574	Db
595	ATGCCAGAGCTGGCTCAGAAAATCAAATGTATTTTGTCTTAGCACCCATAGCCATGTT	654	QY
596		655	QY
574	ATGCCAGAGCTGGCTCAGAAAATCAAATGTATTTTGTCTTAGCACCCATAGCCATGTT	633	Db
575		634	Db
655	AAGCATGCAAAAAGCCCGGACCAAATTTTGTCTGCCAGATATGATCAAGGGA	714	QY
656		715	QY
634	AAGCATGCAAAAAGCCCGGACCAAATTTTGTCTGCCAGATATGATCAAGGGA	693	Db
635		694	Db
715	TTGTTTTGGCAAAAAGAAATTTCTGTATCAGACCAGATTTCTCAGACAACCTTGTTATTATTC	774	QY
716		775	QY
694	TTGTTTTGGCAAAAAGAAATTTCTGTATCAGACCAGATTTCTCAGACAACCTTGTTATTATTC	753	Db
695		754	Db
775	CTTTGTGGCCAGGTGATTTCTGTATCAGATTTGTAGTAATATCATGTTACTTCTGGGTGGA	834	QY
776		835	QY
754	CTTTGTGGCCAGGTGATTTCTGTATCAGATTTGTAGTAATATCATGTTACTTCTGGGTGGA	813	Db
755		814	Db
835	TTCAAACCAACAATATGACAT-----GAGCCGAGCAAGTGTA	873	QY
836		874	QY
814	TTCAAACCAACAATATGACATTTATCTATGGTTTGTTCAGAGCCGAGCAAGTGTA	873	Db
815		874	Db
874	TATGCTGCCCACTCTTGTCTGGAACATCTGTGCAAAATATTTACACTGGAGCCAGGCA	933	QY
875		934	QY
874	TATGCTGCCCACTCTTGTCTGGAACATCTGTGCAAAATATTTACACTGGAGCCAGGCA	933	Db
875		934	Db
934	GTGAATTCCTGGTGAACCTCCGGCAATTTGACTGGGGAGTCAGACCACCAAAATCTGGAATA	993	QY
935		994	QY
934	GTGAATTCCTGGTGAACCTCCGGCAATTTGACTGGGGAGTCAGACCACCAAAATCTGGAATA	993	Db
935		994	Db
994	TGCAATCAGCCAACTCCTGTAAGGTACAGAGTCAGAGATAGACGTCTCCTACAGCAATG	1053	QY
995		1054	QY
994	TGCAATCAGCCAACTCCTGTAAGGTACAGAGTCAGAGATAGACGTCTCCTACAGCAATG	1053	Db
995		1054	Db
1054	TGGACGAGGTACAGGACTGGCTTTCAAATCCAGAAGACGTGAAAATGCTGCTCTCTGAG	1113	QY
1055		1114	QY
1054	TGGACGAGGTACAGGACTGGCTTTCAAATCCAGAAGACGTGAAAATGCTGCTCTCTGAG	1113	Db
1055		1114	Db
1114	GTGACCAACCTCATCTACCAAGAAATATTCCTGAATGGGCTCAGTGGATTTTCATCTGG	1173	QY
1115		1174	QY
1114	GTGACCAACCTCATCTACCAAGAAATATTCCTGAATGGGCTCAGTGGATTTTCATCTGG	1173	Db
1115		1174	Db
1174	GGTTTTGGATGCTCCTCACCCTATCTACAATGAATATCATTCATCTGATGACGAGGAG	1233	QY
1175		1234	QY
1174	GGTTTTGGATGCTCCTCACCCTATCTACAATGAATATCATTCATCTGATGACGAGGAG	1233	Db
1175		1234	Db
1234	ACCAACCTTTCCAGGACCGGTGTGAGCCGCTATTG	1269	QY
1235		1269	QY
1234	ACCAACCTTTCCAGGACCGGTGTGAGCCGCTATTG	1269	Db
1235		1269	Db

RESULT 9	
ID	ABK86570
AC	ABK86570 standard; cDNA; 1116 BP.
XX	
AC	ABK86570;
XX	
DT	24-SEP-2002 (first entry)
XX	
DE	Human cDNA encoding lysosomal acid lipase #2.
XX	
KW	Human; ss; gene; lysosomal acid lipase; lipid malabsorption illness;
KW	cystic fibrosis; alcoholism; heart disease; heart attack;
KW	Wolman disease; cholesterol ester storage disease; brain injury;
KW	mood disorder; anxiety disorder; thought disorder; volition disorder;
KW	sleep disorder; neurogenic disorder; myopathic disorder; COPD;
KW	obesity; cancer; chronic obstructive pulmonary disease; diabetes;
KW	cardiovascular disorder; Alzheimer's disease; Parkinson's disease;
KW	anorexia; osteoarthritis; central nervous system disorder;
KW	peripheral nervous system disorder.
XX	
OS	Homo sapiens.
XX	
FH	Key
FT	Location/Qualifiers
CDS	1..1116
FT	/*tag= a
FT	/product= "Lysosomal acid lipase #2"
XX	
FN	W0200236731-A2.
XX	
PD	10-MAY-2002.
XX	
PF	30-OCT-2001; 2001WO-EP12518.
XX	
PR	31-OCT-2000; 2000US-244170P.
PR	29-MAY-2001; 2001US-293516P.
XX	
PA	(FARB) BAYER AG.
XX	
PI	Xiao Y;
XX	
DR	WPI; 2002-519248/55.
DR	P-PSDB; AAU99165.
XX	
PT	Novel human lysosomal acid lipase polypeptide, useful for treating
PT	cancer, diabetes, obesity, chronic obstructive pulmonary disease,
PT	peripheral or central nervous system disorder or cardiovascular
PT	disorder -
XX	
PS	Claim 1; Fig 4; 126pp; English.
XX	
CC	The invention relates to a purified human lysosomal acid lipase
CC	polypeptide. Also included are the polynucleotide encoding the
CC	lipase (or its fragment, derivative, allele or sequence at least 60%
CC	identical to it), vectors, host cells, a reagent (e.g. an antisense
CC	oligonucleotide) which binds to the lipase or polynucleotide (used for
CC	detection and modulating/reducing the lipase activity) and an anti-l
CC	antibody. The lipase and polynucleotide are useful for identifying
CC	therapeutic agents that either increase or decrease the lipase activ
CC	The identified agent, the lipase and polynucleotide are useful for
CC	treatment of a disease such as lipid malabsorption illness,
CC	cystic fibrosis, alcoholism, heart disease, heart attack,
CC	Wolman disease, cholesterol ester storage disease, brain injury,
CC	mood disorder, anxiety disorder, thought disorder, volition disorder
CC	sleep disorder, neurogenic disorder, myopathic disorder, obesity, ca
CC	chronic obstructive pulmonary disease (COPD), diabetes,
CC	cardiovascular disorder, Alzheimer's disease, Parkinson's disease,
CC	anorexia, osteoarthritis, a central nervous system disorder and
CC	a peripheral nervous system disorder. The present sequence is the
CC	cDNA encoding human lysosomal lipase #2.
XX	
SQ	Sequence 1116 BP; 326 A; 234 C; 245 G; 311 T; 0 other;

Query Match 59.3%; Score 752; DB 24; Length 1116;
 Best Local Similarity 97.4%; Pred. No. 7.5e-229;
 Matches 783; Conservative 0; Mismatches 0; Indels 21; Gaps 1;

QY 148 AGTGAATATCTCAACATCAAGGCTATCCCTGTGAGGAATATGAAGTCGCACTGAAGAT 207
 DB 73 AGTGAATATCTCAACATCAAGGCTATCCCTGTGAGGAATATGAAGTCGCACTGAAGAT 132

QY 208 GGGTATATCTCTTCTGTTAAAGGATTCCTCGAGGCTAGTCAACCTTAAGAGACAGGT 267
 DB 133 GGGTATATCTCTTCTGTTAAAGGATTCCTCGAGGCTAGTCAACCTTAAGAGACAGGT 192

QY 268 TCCAGGCTCTGTGTGTTACTGACGATGCGCTAGTGGAGGTGTAGCAACTGATTTCC 327
 DB 193 TCCAGGCTCTGTGTGTTACTGACGATGCGCTAGTGGAGGTGTAGCAACTGATTTCC 252

QY 328 AACCTGCCCAACATAGCTGGCTTCATCTGCGAGATGCTGCTTTTGACGTGTGATG 387
 DB 253 AACCTGCCCAACATAGCTGGCTTCATCTGCGAGATGCTGCTTTTGACGTGTGATG 312

QY 388 GGGAAACAGCAGGGGAAACCGCTGTCTCGAAGACACAGACACTCTCCATAGACCAAGAT 447
 DB 313 GGGAAACAGCAGGGGAAACCGCTGTCTCGAAGACACAGACACTCTCCATAGACCAAGAT 372

QY 448 GAGTCTGGGCTTCAGTTATGATGATGCTAGTGTGACCTTCTGCTGAGTGAATAC 507
 DB 373 GAGTCTGGGCTTCAGTTATGATGATGCTAGTGTGACCTTCTGCTGAGTGAATAC 432

QY 508 TTTATTTTGCAGAAAACCGGCGAGGAAAGATCTATTATGTGCGCTATTCACAGGGCACC 567
 DB 433 TTTATTTTGCAGAAAACCGGCGAGGAAAGATCTATTATGTGCGCTATTCACAGGGCACC 492

QY 568 ACCATGGCTTTATGCAATTTTCCACCTGCGAGCTGGCTCAGAAAATCAAAATGTAT 627
 DB 493 ACCATGGCTTTATGCAATTTTCCACCTGCGAGCTGGCTCAGAAAATCAAAATGTAT 552

QY 628 TTTGCTTTAGCACCACATAGCACTGTTAAGCATGCAAAAAGCCCGGACCAAAATTTTG 687
 DB 553 TTTGCTTTAGCACCACATAGCACTGTTAAGCATGCAAAAAGCCCGGACCAAAATTTTG 612

QY 688 TTGCTGCCAGATATGATGATCAAGGATGTTTGGCAAAAAGAAATTTCTGTATCAGACC 747
 DB 613 TTGCTGCCAGATATGATGATCAAGGATGTTTGGCAAAAAGAAATTTCTGTATCAGACC 672

QY 748 AGATTTCTCAGCACTGTTATTTACCTTTGTCGAGGATGTTCTGTATCAGATTTGT 807
 DB 673 AGATTTCTCAGCACTGTTATTTACCTTTGTCGAGGATGTTCTGTATCAGATTTGT 732

QY 808 AGTAATATCATGTTACTTCTGGGTGATTTCAACACCAACCAATATGAACAT 857
 DB 733 AGTAATATCATGTTACTTCTGGGTGATTTCAACACCAACCAATATGAACAT 792

QY 858 -----GAGCGGACGAGTGTATATGCTGCCACACACTCTGCTGGACATCTGTG 906
 DB 793 GGTGTTGTACAGACCGAGCAAGTGTATATGCTGCCACACACTCTGCTGGACATCTGTG 852

QY 907 CAAATAATTTCTACACTGGAGCCAG 930
 DB 853 CAAATAATTTCTACACTGGAGCCAG 876

RESULT 10
 ABN97383
 ID ABN97383 standard; DNA; 2481 BP.
 AC ABN97383;
 XX
 DT 13-AUG-2002 (first entry)
 DE Gene #3881 used to diagnose liver cancer.
 DE Gene; liver cancer; ds; hepatocellular carcinoma; hepatotropic;
 KW metastatic liver tumour; cytostatic; expression profile; disease state;
 KW

disease progression; drug toxicity; drug efficacy; drug metabolism.
 OS Homo sapiens.
 XX WO200229103-A2.
 XX 11-APR-2002.
 XX 02-OCT-2001; 2001WO-US30589.
 XX 02-OCT-2000; 2000US-237054P.
 XX (GENE-) GENE LOGIC INC.
 PA Horne D, Alvares C, Peres-Da-Silva S, Vockley JG;
 XX WPI; 2002-426119/45.
 XX
 XX Diagnosing and detecting the progression of liver cancer,
 PT hepatocellular carcinoma or metastatic liver tumor in a patient,
 PT involves detecting the level of expression of two or more genes in a
 PT liver tissue sample -
 XX
 PS Claim 1; SEQ ID NO 3881; 298pp; English.
 XX
 CC The invention relates to a novel method for diagnosing and detecting the
 CC progression of liver cancer, hepatocellular carcinoma or metastatic liver
 CC tumour in a patient, and differentiating metastatic liver cancer from
 CC hepatocellular carcinoma in a patient, involving detecting the level of
 CC expression of two or more genes represented in ABN93503-ABN97455 in a
 CC tissue sample. The method of the invention has hepatotropic, and
 CC cytostatic activity. The method is useful for diagnosing and detecting
 CC the progression of liver cancer, hepatocellular carcinoma and metastatic
 CC liver carcinoma in a patient. The method is useful for identifying
 CC expression profiles which serve as useful diagnostic markers as well as
 CC markers that can be used to monitor disease states, disease progression,
 CC drug toxicity, drug efficacy and drug metabolism.
 CC Note: The sequence data for this patent did not form part of the printed
 CC specification, but was obtained in electronic format directly from WIPO
 CC at fcp.wipo.int/pub/published_pct_sequences.

Sequence 2481 BP; 692 A; 477 C; 513 G; 799 T; 0 other;
 Query Match 41.6%; Score 527.8; DB 24; Length 2481;
 Best Local Similarity 66.6%; Pred. No. 4e-157;
 Matches 751; Conservative 0; Mismatches 376; Indels 0; Gaps 0;

QY 115 AAAGCTGTGAGCCCAAGCACTTATGATGATGATGATGATGATGATGATGATGATGATGAT 174
 DB 112 ACAGCTGTGATCTCTGAAACCAACATGATGATGATGATGATGATGATGATGATGATGATGAT 171

QY 175 CCTGTGTGAGGAATATGAGTGTGCACTGAGATGAGTGTGATGATGATGATGATGATGATGATGAT 234
 DB 172 CCTGTGTGAGGAATATGAGTGTGCACTGAGATGAGTGTGATGATGATGATGATGATGATGATGAT 231

QY 235 CCTGTGTGAGGAATATGAGTGTGCACTGAGATGAGTGTGATGATGATGATGATGATGATGATGAT 294
 DB 232 CCTGTGTGAGGAATATGAGTGTGCACTGAGATGAGTGTGATGATGATGATGATGATGATGATGAT 291

QY 295 GGCCTAGTGTGAGGATGCTAGCACTGATGATGATGATGATGATGATGATGATGATGATGATGAT 354
 DB 292 GGCCTAGTGTGAGGATGCTAGCACTGATGATGATGATGATGATGATGATGATGATGATGATGAT 351

QY 355 ATTTCTGGCAGATGCTGTTTTCACCTGTGATGATGATGATGATGATGATGATGATGATGATGATGAT 414
 DB 352 ATTTCTGGCAGATGCTGTTTTCACCTGTGATGATGATGATGATGATGATGATGATGATGATGATGAT 411

QY 415 CGAAAAACACAGACACTCTCCATAGACGATGATGATGATGATGATGATGATGATGATGATGATGAT 474
 DB 412 CGAAAAACACAGACACTCTCCATAGACGATGATGATGATGATGATGATGATGATGATGATGATGAT 471

QY 475 ATGGCTAGGTTTTCACCTTCTGCTGAGTGAATACTTTATTTTTCAGAAAAACGGCCAGGAA 534

472 ATGGCAAAATATGACCTACACCTGCTCCATTAACCTTCTGAAATAAAACGCGCCAAAGAA 531
 535 AAGATCTATTATGCTGGCTATTACAGGGCACCACCATGGCTTTATTCATTTCCACC 594
 532 CAAGTGATTATGTTGGTCACTTCTCAAGGCACCATATAGTTTATAGCATTTTACAG 591
 595 ATGCCAGAGCTGGCTCAGAAAATCAAAATGATTTTGTCTTTAGCACCACCATAGCCATGTT 654
 592 ATCCCTGAGCTGGCTAAAGGATTAAGATGTTTTTGGCCCTGGTCTGCTGGCTCCGTC 651
 655 AAGCATCAAAAGCCCGGACCAAAATTTTGTCTGTCAGATATGATGATCAAGGA 714
 652 GCCTTCTGTACTAGCCCTATGGCCAAATTAGGACGATTAACAGATCATCTCAATAGGAC 711
 715 TTGTTTGGCAAAAGAAATTTCTGTATCAGACCAGATTTCTCAGACAACTTGTATTATAC 774
 712 TTATTTGGACAAAGAAATTTCTCCAGAGTGGTTTTTGAAGTGGCTGGGTACCCAC 771
 775 CTTTGTGGCAGAGTGATCTTGTATCAGATTTGTAGTAATATCATGTTACTTCTGGGTGA 834
 772 GTTTGCACTCATGTCTACTGAAAGGAGCTCTGTGGAATCTCTGTTTCTCTGTGTGA 831
 835 TTCAACACCAACATATGAACATGAGCGGAGCAAGTGTATATGCTGCCACACTCTTGT 894
 832 TTTATGAGAGAAATTTAAATATGTTCTAGAGTGGATGATATACACACATTTCTCTGCT 891
 895 GGAACATCTGTGCAAAATATTTCTACACTGGAGCCAGCAGTGAATTTCTGGTAACTCCG 954
 892 GGAACATCTGTGCAAAATATTTCTACACTGGAGCCAGCTGTTAAATTTCCAAAGTTTCAA 951
 955 GCATTTGACCTGGGGAGTGAGACCAAAATCTGMAAAATGCAATCAGCCAACTCTGTGA 1014
 952 GCCTTTGACCTGGGAAGCAGTGCCCAAGATTTATTTCAATTAACACAGATTAATCTCC 1011
 1015 AGGTACAGAGTCAGAGATATGACCGTCTCTACAGCAATGTGGACAGAGTCAAGACTGG 1074
 1012 ACATACATGTGAAGGACATGCTTGTGCCACTGCACTCTGGAGCGGGGTCAAGACTGG 1071
 1075 CTTTCAATCCAGAGAGCGTGAATGCTGCTCTGAGGTGACCACTCATCTACCAT 1134
 1072 CTTGAGATGCTACGAGCTCAATATCTTACTGACTCAGATCACTCACTTGGTGTCCAT 1131
 1135 AAGAAATATCTGAAATGGGCTCAGTGGATTTCTATCTGGGGTTTGGATGCTCCCTACCGT 1194
 1132 GAGAGCATTCGGAAATGGAGCATTTGACTTCAATTTGGGGCTGGATGCCCTTGGAGG 1191
 1195 ATGTACAAATGAATCATCTCATGATGAGCAGAGGAGGAGCAACT 1241
 1192 CTTTATAATAAAATTTAATCTAATGAGGAAATATCAGTGAAGCT 1238

RESULT 11

AAH02906 standard; DNA; 2493 BP.

XX AAH02906;

XX 15-JUN-2001 (first entry)

XX Human shear stress-response coding sequence SEQ ID NO: 65.

XX Human; shear stress-response protein; vascular disease;

XX arteriosclerosis; ds.

XX Homo sapiens.

XX WO200125427-A1.

XX 12-APR-2001.

XX 02-OCT-2000; 2000WO-JP06840.

XX 01-OCT-1999; 99JP-0280976.

XX (KYOW) KYOWA HAKKO KOGYO KK.
 PA (NOJI/) NOJIMA H.
 XX Nojima H, Yoshisue H, Obayashi M, Ota T, Kawabata A, Sakurada K;
 PI Kuga T, Sekine S, Nakamura Y, Sugano S;
 XX WPI; 2001-266308/27.
 DR P-PSDB; AAB90783.
 XX DNA sequences, proteins encoded by them and antibodies against them
 PT useful in diagnosis and treatment of vascular disease caused by
 PT arteriosclerosis -
 XX Claim 20; Page 398-402; 678pp; Japanese.
 CC The present invention provides the protein and coding sequences of a
 CC number of human shear stress response proteins. These are useful in the
 CC diagnosis, treatment and screening of vascular diseases caused by
 CC arteriosclerosis, including heart failure, post-PTCA restenosis and
 CC hypertension.
 XX Sequence 2493 BP; 692 A; 484 C; 516 G; 801 T; 0 other;

Query Match 41.5%; Score 526.2; DB 22; Length 2493;

Best Local Similarity 66.6%; Pred. No. 1.3e-156; Indels 0; Gaps 0;

Matches 751; Conservative 0; Mismatches 376;

QY 115 AAAGCTGTGGACCCAGAGCAATTCATGAATATTAGTGAATATCAATCAATCAAGGTAT 174

Db 119 ACAGCTGTGGATCTGAAACAAACATGATGTGAGTGAATATCTCTTACTGGGATTC 178

QY 175 CCCTGTGAGGAATATGAGTCGCACTGAGATGGGTATATCTTCTTGTAAACAGATT 234

Db 179 CCTAGTGGGAATACCTAGTTGACAGAGATGGATATATCTGTGCTTAAACCAATT 238

QY 235 CCTCAGGCCCTAGTGCACCTTAAGAGACAGAGTTTCCAGGCTGTGGTGTACTGACGAT 294

Db 239 CCTCATGGGAGGAAGAACCACTCTGACAAAGTCCCAACAGTTGTCTTCTGCAACAT 298

QY 295 GGCCTAGTTGAGGTGTAGCACTGGATTTCCAACTGCCCAACATAGCTGGGCTTC 354

Db 299 GCTTGTCTGGCAGATTTCTAGTAATCTGGTCAAACTTGCACACAGAGCTGGGCTTC 358

QY 355 ATCTGGCAGATGCTGGTTTTGAGCTGTGGATGGGAAACAGCAGGGGAAACGCTGTCT 414

Db 359 ATCTGTGTATGCTGGTTTTGAGCTGTGGATGGGAAACAGCAGGGGAAATCTGTCT 418

QY 415 CGAAACACAAAGACACTCTCCATAGA CCAAGATGAGTTCTGGGCTTTTCACTTATGATGAG 474

Db 419 CGGAAACATAAGACACTCTCAGTTTCTCAGGATGAATTTCTGGGCTTTTCACTTATGATGAG 478

QY 475 ATGGCTAGGTTTGGCTTCCCTGAGTGAATATTTTGTGCAAGAAACGGGCGCAGAA 534

Db 479 ATGGCAAAATATGACCTTACAGCTTCCATTTACTTCAATTTGAATAAATCTGGCCAGAA 538

QY 535 AAGATCTATTATGCTGGCTATTTCACAGGCAACCACTGGGCTTTTATGCAATTTCCACC 594

Db 539 CAAGTGTATTATGTTGGTCACTTCTCAAGGCACCACTATAGGTTTATAGCATTTTACAG 598

QY 595 ATGCCAGAGCTGGCTCAGAAAATCAAAATGATTTTGTCTTTAGCACCACCATAGCCATGTT 654

Db 599 ATCCCTGAGCTGGCTAAAGGATTAAGATGTTTTTGGCCCTGGGCTCTGTGGCTCCGTC 658

QY 655 AAGCATCAAAAGCCCGGACCAAAATTTTGTGCTGTCAGATATGATGATCAAGGA 714

Db 659 GCCTTCTGTACTAGCCCTATGGCCAAATTAGGACGATTTACAGATCATCTCATTAAGGAC 718

QY 715 TTGTTGGCAAAAGAAATTTCTGTATCAGACCAAGATTTCTCAGACAACTTGTATTATAC 774

Db 719 TTTATTTGGAGACAAAGAAATTTCTTCCCGAGAGTGGTTTTTGAAGTGGCTGGGTACCCAC 778

QY 775 CTTTGTGGCCAGGTGATCTTGTGATGATGATTTGTAGTATATATCATGTTTACTTCTGGGTGA 834

Db 779 GTTTCACCTCATGTCTATGCTAGGAGCTCTGTGAAATCTCTGTCTCTGTGGA 838
 Qy 835 TTCAACACCAACAATATGAATGAGCGAGCAAGTGATATGCTGCCACACCTCTTGCT 894
 Db 839 TTTATGAGAGAAATTTAAATATGCTAGAGTGATATGATACAAACATCTCTGCT 898
 Qy 895 GGAACATCTGTGCAAAATATTTCTACATGGAGCCAGGAGTGAATTTCTGTGAACTCGG 954
 Db 899 GGAACCTTCTGTGCAAAACATGTTACATGGAGCCAGGCTGTTAAATTCCAAAGTTTCAA 958
 Qy 955 GCATTTGACTGGGAGTGAGACCAAAATCTGGAATAATGCAATCAGCCAACTCTGTA 1014
 Db 959 GCTTTGACTGGGAGGAGTGAGCCAGAAATTTTTCATTACACAGAGTTA-TCTCTCC 1018
 Qy 1015 AGGTACAGAGTCAGAGATATGAGCTGCTCTACAGCAATGTGACAGAGGTCAGGACTGG 1074
 Db 1019 ACATACAAATGTGAGGACATGCTTGTGCGAGCTGAGCTGAGCGGGGTACGAGCTGG 1078
 Qy 1075 CTTTCAATCCAGAGAGTGAAATGCTGCTCTCTGAGGTGACCAACTCATCTACCAT 1134
 Db 1079 CTTGAGATGTCTAGGAGTCAATATCTTACTGACTCAGATCACCACCTTGGTGTCCAT 1138
 Qy 1135 AAGAATATTTCTGAATGGCTCAGCTGAGTTTCACTGCGGTTTGGATGCTCTCACCGT 1194
 Db 1139 GAGAGCATTCGGATGGAGCATCTTGACTTCACTTTGGGCTGGATGCTCTGGAGG 1198
 Qy 1195 ATGTCAATGAAATCATCATCTGATGAGCGAGGAGAGACCAACT 1241
 Db 1199 CTTTATAATAAATTAATTAATCTAATGAGAAATATCAGTGAAGCT 1245

RESULT 12

AAQ42310
 ID AAQ42310 standard; DNA; 1378 BP.
 AC AAQ42310;
 DT 25-MAR-2003 (updated)
 DT 20-SEP-1993 (first entry)
 XX
 DE RGL precursor.
 XX
 KW Rabbit gastric lipase; RGL; prGLN2.1; fat; bioconversion;
 KW hydrolysis; transesterification; ds.
 XX
 OS Oryctolagus cuniculus.
 XX
 FH Key Location/Qualifiers
 FT misc_RNA 120...1247
 FT /tag= a
 FT /note= "claim 1; page 10-11"
 XX
 PN EP542629-A1.
 XX
 PD 19-MAY-1993.
 XX
 PF 12-NOV-1992; 92EP-0403055.
 XX
 PR 13-NOV-1991; 91FR-0013948.
 XX
 PA (LJOU) INST RECH JOUVEINAL.
 XX
 PI Benicourt C, Blanchard C, Junien J;
 XX
 DR WPI; 1993-161080/20.
 DR P-PSDB; AAR37302.
 XX
 PT Rabbit gastric lipase, its precursor and their DNA - useful for
 PT treating conditions linked to gastric lipase deficiency, such as
 PT mucoviscidiosis and pancreatic exocrine insufficiency
 XX
 PS Claim 1; Fig 6; 31pp; French.

XX mRNA was isolated from rabbit gastric mucosal cells, converted to
 CC cDNA and fragments cloned in pUC18. Recombinant plasmids were used
 CC to transform E. coli MW294 and cells screened for reactivity with
 CC probe corresp. to known parts of the RGL sequence. One positive
 CC clone contd. plasmid pJ0101 which contd. a 1.35 kb PstI sequence.
 CC This fragment was isolated, subjected to PCR and the amplification
 CC prod. ligated with the 2.2 kb BglII-SacI fragment of prU276 (contg.
 CC a synthetic Tac promoter and transcription stop signal). The
 CC recombinant plasmid was designated prGLN2.1.
 CC (Updated on 25-MAR-2003 to correct PN field.)
 XX
 SQ Sequence 1378 BP; 397 A; 287 C; 276 G; 418 T; 0 other;
 Query March 38.6%; Score 489.2; DB 14; Length 1378;
 Best Local Similarity 65.0%; Pred. No. 5.8e-145;
 Matches 722; Conservative 0; Mismatches 388; Indels 0; Gaps 0;
 Qy 125 ACCGAGAGCATTTCATGATATATAGTGAATATCATCAACATCAAGCTATCCCTGTGAGG 184
 Db 136 ACCCTGAAAGTGAATATGATATATAGTCAGATGATTTCTTACTGGGATACCCAAGTGAA 195
 Qy 185 AATATGAAGTCGCAACTGAAGATGGGTATATCTTCTGTAAACAGAGTTCCTCTAGGCGC 244
 Db 196 AATATGAAGTTGTGACTGAAGATGGTATATCTTGAAGTTAACAGAAATCTCTATGGGA 255
 Qy 245 TAGTGCACCTAAGAGAGACAGGTTCCAGGCTGTGGTGTACTGACAGCATGGCTTAGTTG 304
 Db 256 AGAAAAATTCAGGGAACAGAGACCCGCTGTATTTTTCAGCATGGTTGCTTG 315
 Qy 305 GAGGTGCTAGCACTGGATTTCCAACTGCCCAACAAATAGCTGGGCTTCATTCTGGCAG 364
 Db 316 CATCAGCCTCAAACTGGATCTCCACTGCCCAACAGAGCTGGCTTTATCTTGGCAG 375
 Qy 365 ATGCTGGTTTGTGAGTGTGGTGAACAGAGGGGAAACCGCTGTCTCGAAACACA 424
 Db 376 ATGCTGGTTATGTTGTGGCTGGGGAACAGAGAGGAAATACCTGGTCCAGAGAAACT 435
 Qy 425 AGACACTCTCCATAGACCAAGATGAGTCTGGGCTTTTCAAGTTATGATGAGTGGTAGT 484
 Db 436 TATACATTTCCAGAGCTCGTTGAATCTGGGCTTTTGTAGCTTTGATGAATAGGCTAAAT 495
 Qy 485 TTGACCTTCTCTGAGTGATAAACTTTATTTTGCAGAAAAACGGGCCAGGAAAAAGATCTATT 544
 Db 496 ATGACCTTCCAGCCACCAATGTGACTTCATTTGTAAGGAAACTGGACAGGAAAAGTCACT 555
 Qy 545 ATGTCGGCTATTTCAGAGGACCCACCATGGCTTTATTTGCAATTTTCCACCATGCCAGAGC 604
 Db 556 ATGTTGGTCAATTCAGGAGCCACCATTTGGTTTATTTGCTCTTCTCAAAATCCCAAGC 615
 Qy 605 TGGCTCAGAAAAATCAAAATGTATTTTGTCTTAGCACCCTAGCCTAGTGTAAAGCATGCAA 664
 Db 616 TGGCTGAAGAAATCAAAACCTTTTATGATGATGATGATGATGATGATGATGATGATGAT 675
 Qy 665 AAAGCCCCGGGACCAAAATTTTGTGCTGCCAGATATGATGATGATGATGATGATGATGATG 724
 Db 676 AAAGCCTTGTAAACAAACTTAGGTTTATTTCTCCCAACCTGTTTCAAGATATATTTGGTG 735
 Qy 725 AAAAGAAATTTCTGATCAGATGATGATGATGATGATGATGATGATGATGATGATGATG 784
 Db 736 ACAAATATTTCTACCCACACAAATTTCTTGTATCAATTTCTGCCACTCAAGTGTGTCC 795
 Qy 785 AGGTGATTTCTTGTATCAGATTTGTAGTAAATATATGTTACTTCTGGGTGGAATCAACACA 844
 Db 796 GTGAGACACTGAAATGTAATTTGAGCAATGCTTTATTTATCATTTTGTGGCTTTGACAGG 855
 Qy 845 ACATATGAACATGAGCCGAGAGTGTATATGCTGCCACACTCTTGTCTGGAACATCTG 904
 Db 856 CAAACTTGAACATGAGTGGCTTGGATGTGTAGCTGTGATGATGATGATGATGATGATGATG 915
 Qy 905 TGCAAAATATTTTACACTGAGCGAGGAGTGAATTTCTGTGTAATCTCCGGGCAATTTGACT 964
 Db 916 TTCAAAACATGCTGCACTGGACCCAGGCTGTTAAATCTGAAATTTTCAAGCTTTTAAAT 975

QY 1139 ATATTCTGATGGGCTACGTGGATTTCATCTGGGGTTGGATGCTCTCCACCGTATGT 1198
 Db 1143 AGATTCTCTTTTACAACTCACTTGGACTTTATCTGGCAATGATGCCCTCAGAGATT 1202

QY 1199 ACAATGAATCATCCATCTGAT 1220
 Db 1203 ACAATGACATTTGTTCTATGAT 1224

RESULT 14

AAN60566
 ID AAN60566 standard; DNA; 1367 BP.

AC AAN60566;
 XX
 XX 25-MAR-2003 (updated)
 DT 22-AUG-1991 (first entry)
 XX

Sequence encoding human pregastric lipase.

XX Cystic fibrosis therapy; enzyme; lipase deficiency; ss.

XX Homo sapiens.

XX Key Location/Qualifiers
 FH CDS 47..103
 FT /*tag= a
 FT mat_peptide 104..1243
 FT /*tag= b

XX WO8601532-A.

XX 13-MAR-1986.

XX 15-AUG-1985; 85WO-GB00364.

XX 21-AUG-1984; 84GB-0021210.

XX 15-AUG-1985; 85WO-GB00364.

XX 01-JAN-1986; 86GB-0008897.

XX (CLLT) CELLTech LTD.

XX (LOWE/) LOWE P A.

XX Lowe PA;

XX WPI; 1986-081634/12.

XX P-PSDB; AAP60658.

XX New gastric lipase protein, esp. of human origin - for treating

XX lipase deficiency, and DNA sequences coding for it

XX Disclosure; Fig 3; 39pp; English.

XX The inventors claim a pregastric lipase protein and a gene encoding

XX it. Gastric lipase protein is useful for oral administration to

XX treat lipase deficiency, e.g. cystic fibrosis or pancreatitis.

XX (Updated on 25-MAR-2003 to correct PA field.)

XX Sequence 1367 BP; 397 A; 284 C; 269 G; 417 T; 0 other;

XX Query Match 37.0%; Score 470; DB 7; Length 1367;

XX Best Local Similarity 64.2%; Pred. No. 7.7e-139;

XX Matches 707; Conservative 0; Mismatches 395; Indels 0; Gaps 0;

QY 119 CTGTGGACCCAGAGCAATTCATGAATATTAGTGAATATCCACATCAAGGCTATCCCT 178

Db 123 CTGGAAGCCCTGAAGTGACTATGAACATTAGTCAGATGATTACTTATTGGGGATACCCAA 182

QY 179 GTGAGGAATATGAAGTCGCAACTGAAGATGGGTATATCTCTGTTAAACAGATTTCCTC 238

Db 183 ATGAGAAATATGAAGTTGTGACTGAGAGATGGTATTATCTTTGAAGTCAATAGAAATTCCTT 242

XX

QY 239 GAGGCTTAGTCACCACTAAGACAGACAGGTTCCAGGCTGTGGTGTACTGCAGCATGGCC 298
 Db 243 ATGGGAAGAAAAATTCAGGGAATACAGGCCAGAGACTGTGTGTGTGTGTGTGTGTGTGT 302
 QY 299 TAGTTGAGAGTGTAGCACTGATTTCCAACTGCTGCCACATAGACTGCTGGGCTTATTC 358
 Db 303 TGCTTGATCAGGCCACAACTGATTTCCAACTGCTGCCACATAGACTGCTGGGCTTATTC 362
 QY 359 TGGCAGATGCTGGTGTGTGACGTGTGATGGGGAACAGCAGGGGAAACGCCGTGCTCTCGAA 418
 Db 363 TGGCAGATGCTGGTGTGTGATGATGTGTGGCTGGCAACAGCAGAGGAACACCTGGGCCAGAA 422
 QY 419 AACAAAGACACTCTCCATAGACCAAGATGAGTCTGGGCTTTTCAGTTATGATGATGG 478
 Db 423 GAAACTTTGACTATTACACAGATTTCAGTTGAAATCTGGGCTTTTCAGTTGATGAAATGG 482
 QY 479 CTAGGTTTGACCTTCTCTGAGTGTATAAACTTTTATTTTGCAGAAAAACGGGACAGAAAAAGA 538
 Db 483 CTAAATATGACCTTCCAGCCACCAATCGACTTCAATTGTAAGAAAACTGGACAGACGAGC 542
 QY 539 TCTATTATGTCGGCTATTTCACAGGCGACCACTATGGGCTTTTATTCATTTTCCACCATGC 598
 Db 543 TACACTATGTTGGCCATTCCAGGCGACCACTATGGTTTTATTTGCTTTTCCACCAATC 602
 QY 599 CAGAGCTGGCTCAGAAATCAAAATGATTTTCTTGTAGCACCACCATAGCACTGTTAAGC 658
 Db 603 CCAGCTGGCTTAAAGAAATCAAAACCTTCTATGCTTAGCTTCTGTTGCTGCTGCTGCTGCT 662
 QY 659 ATGCAAAAAGCCCGGGACCAAAATTTTGTGTGTCAGATATGATGATCAAGGATTTGT 718
 Db 663 ATACAAAAGCCCTATATAACCAAACTTAGATTGCTTCTCAATCCCTCTTCAAGTTATAT 722
 QY 719 TTGGCAAAAAGAAATTTCTGTATCAGACAGATTCTCAGACAACTGTTGTTTATTTACCTTT 778
 Db 723 TTGGTGACAAAATATTCTACCCACCAACTTCTTGATCAATTTCTTGCTACTGAAATGT 782
 QY 779 GTGGCCAGGTGATTTCTTGATCAGATTGTTAGTAAATATCATGTTACTTCTGGGTGGATTCA 838
 Db 783 GCTCCCGTGAGTGTGAATCTCTTTGAGCAATGCTTATTATAATTTGTTGATTTG 842
 QY 839 ACACCAACAATATGAACATCAGCGCAGCAAGTGTATATGTCGCCACACACTCTTTGCTGGAA 898
 Db 843 ACAGTAAGAACTTTAACACAGATCGCTTGGATGTATCTATCACATAATCCAGCAGGAA 902
 QY 899 CATCTGTGCAAAATATTCTACACTGGAGCCAGCAGTGAATCTGTTGTTGTTGTTGTTGTTGTT 958
 Db 903 CTCTCTGTTCAAAACATGTTCCATTGGACCCAGGCTGTAAAGTCTGGGAAATTCCTAAGCTT 962
 QY 959 TTGACTGGGGAGTGAGACCAAAAATCTGAAAAAATGCAATCAGCAAACTCTCTGTAAAGT 1018
 Db 963 ATGACTGGGAGGCCAGTTCCAGATAGGATGCACTATGATGATGATGATGATGATGATGATGAT 1022
 QY 1019 ACAGAGTCAGAGATATGACGGTCCCTACAGCAATGTGGACAGAGGTGAGGATGAGGCTTT 1078
 Db 1023 ACAATGTGACAGCCATGAATGTACCAATTCAGTTGTGGAAACGGTGGCAAGGACCTGTTGG 1082
 QY 1079 CAATCCAGAAAGCGTGAATAATGCTGCTCTCTGAGGTGACCAACCTCATCTACCAATAAGA 1138
 Db 1083 CTGACCCCAAGATGTTGGGCTTTTGCTTCCAAAACCTCCCAATCTTATTATCCACAGG 1142
 QY 1139 ATATTCTGAAATGGGCTCAGCTGGATTTCACTGGGCTTTGGATGCTCTCAACCGTATGT 1198
 Db 1143 AGATTCTCTTTTAACTAATCACTTGGACTTATCTATGGGCAATGGATGCTCTCAAGAAAGTTT 1202
 QY 1199 ACAATGAATCATCCATCTGAT 1220
 Db 1203 ACAATGACATTTTCTATGAT 1224

RESULT 15

AAT58916

ID AAT58916 standard; DNA; 1367 BP.

XX

AC AAT58916;
 XX 19-AUG-1997 (first entry)
 DT Human gastric lipase coding sequence.
 XX
 KW Duodenal; gastric; lipase; transgenic; plant; recombinant; extract; ss;
 KW food; absorption; fat; pancreatic; cystic fibrosis; exocrine; dairy;
 KW hydrolysis; trans-esterification; substrate; enzyme; biofuel; human.
 XX
 OS Homo sapiens.
 XX
 FH Key Location/Qualifiers
 FT CDS /..1243
 FT /*tag= a
 FT /product= gastric lipase
 FT sig_peptide 47..103
 FT /*tag= b
 FT mat_peptide 104..1240
 FT /*tag= c
 XX
 FN WO9633277-A2.
 XX
 XX 24-OCT-1996.
 XX
 XX 19-APR-1996; 96WO-FR00606.
 XX
 XX 20-APR-1995; 95FR-0004754.
 XX
 XX (BIOC-) BIOCEM SA.
 XX (LJOU) INST RECH JOUVEINAL.
 XX
 XX Baudino S, Benicourt C, Cudrey C, Gruber V, Lenee P;
 XX Merot B;
 XX
 XX WPI; 1996-485783/48.
 XX P-ESDB; AAW09383.
 XX
 XX Recombinant human or canine pre-duodenal lipase prodn. in transgenic
 XX plants - useful for facilitating absorption of fat, as bio-catalysts
 XX and for prodn. of bio-fuel
 XX
 XX Claim 3; Fig 4; 130pp; French.
 XX
 XX This is the nucleotide sequence encoding the human pre-duodenal (i.e.
 XX gastric) lipase enzyme. The sequence can be used to generate transgenic
 XX plants producing recombinant lipase in an enzymatically active form.
 XX Alternatively bases encoding amino acids 20-23 or 20-73 (1-4 or 1-54 of
 XX the mature protein, respectively) can be deleted to form the derivatives
 XX designated Delta-4 or Delta-54 respectively. Plants, or their extracts,
 XX expressing the lipases or the truncated derivatives, can be used:
 XX (a) as pharmaceuticals or food to facilitate absorption of fat, either
 XX in healthy subjects or in patients with inadequate levels of
 XX gastric/pancreatic lipase e.g. those with cystic fibrosis or exocrine
 XX pancreatic insufficiency, in the elderly or in patients undergoing medical
 XX treatment which alters fat absorption; (b) for performing industrial or
 XX agricultural reactions, e.g. in processing of fats or in the dairy
 XX industry, for hydrolysis or transesterification reactions, etc., where
 XX the plant material may provide both enzyme and substrate. The transgenic
 XX plants can also be used for biofuel production.
 XX
 XX Sequence 1367 BP; 397 A; 284 C; 269 G; 417 T; 0 other;
 XX
 XX Query Match 37.0%; Score 470; DB 17; Length 1367;
 XX Best Local Similarity 64.2%; Pred. No. 7.7e-139;
 XX Matches 707; Conservative 0; Mismatches 395; Indels 0; Gaps 0;
 XX
 XX 119 CTGTGGACCCGAGCATTCTGATATATGTAATCATCCATCAAGGCTATCCCT 178
 XX |||||
 XX 123 CTGGAAGCCCTGAAGTACTATGACATTAATGATGATTTATTTGGGATACCAA 182
 XX |||||
 XX 179 GTGAGGAATATGAATCGCACTGAAGATGGGTATATCTTCTGTTAAACAGAGATTCCTC 238
 XX |||||

Search completed: February 19, 2004, 05:06:03
 Job time : 356.095 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2004 Compugen Ltd.

OM nucleic - nucleic search, using sw model

Run on: February 19, 2004, 04:49:02, Search time 78.5225 Seconds
(without alignments)
7133.182 Million cell updates/sec

Title: US-10-042-431-46
Perfect score: 1369
Sequence: 1 atgttggaacctgtcaag.....gacggtgtgagccgtattg 1369

Scoring table: IDENTITY NUC
Gapop 10.0, Gapext 1.0

Searched: 569978 seqs, 220691566 residues

Total number of hits satisfying chosen parameters: 1139956

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database: Issued Patents NA.*

- 1: /cgm2_6/ptodata/2/ina/5A_COMB.seq.*
- 2: /cgm2_6/ptodata/2/ina/5B_COMB.seq.*
- 3: /cgm2_6/ptodata/2/ina/6A_COMB.seq.*
- 4: /cgm2_6/ptodata/2/ina/6B_COMB.seq.*
- 5: /cgm2_6/ptodata/2/ina/PCTUS_COMB.seq.*
- 6: /cgm2_6/ptodata/2/ina/backfile1.seq.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	470	37.0	1367	4	US-09-348-930A-8
2	456.8	36.0	1411	4	US-09-186-489-3
3	456.8	36.0	1411	4	US-10-043-665B-3
4	453.6	35.7	1134	4	US-09-186-489-4
5	453.6	35.7	1134	4	US-10-043-665B-4
6	450.4	35.5	1134	4	US-09-186-489-1
7	450.4	35.5	1134	4	US-10-043-665B-1
8	449.8	35.4	8324	4	US-09-186-489-5
9	449.8	35.4	8324	4	US-10-043-665B-5
10	444.2	35.0	1137	1	US-08-227-108-2
11	444.2	35.0	1137	2	US-09-073-674-2
12	444.2	35.0	1140	1	US-08-227-108-4
13	444.2	35.0	1140	2	US-09-073-674-4
14	444.2	35.0	1146	1	US-08-227-108-6
15	444.2	35.0	1146	2	US-09-073-674-6
16	444.2	35.0	1198	4	US-09-348-930A-5
17	444.2	35.0	1528	2	US-08-227-108-1
18	444.2	35.0	1528	2	US-09-073-674-1
19	444.2	35.0	1528	4	US-09-348-930A-1
20	444.2	35.0	1528	4	US-09-348-930A-7
21	414.2	32.6	1360	4	US-09-820-001-1
22	388.8	30.6	1048	4	US-09-348-930A-3
23	104.2	8.2	22067	4	US-09-820-001-3
24	48.2	3.8	178	1	US-08-751-782-2
25	48.2	3.8	178	2	US-08-925-171-2
26	46.4	3.7	43676	3	US-09-356-952-12
27	38.6	3.0	1664976	4	US-08-916-421B-1

28	36.6	2.9	3271	2	US-08-852-806-1	Sequence 1, Appli
29	36.6	2.9	3271	3	US-09-163-669-1	Sequence 1, Appli
30	36.6	2.9	3282	1	US-08-276-852-154	Sequence 154, App
31	36.6	2.9	3282	1	US-08-276-852-154	Sequence 154, App
32	36.6	2.9	3282	1	US-08-899-575-154	Sequence 154, App
33	36.6	2.9	3282	1	US-08-899-575-154	Sequence 154, App
34	36.6	2.9	3282	1	US-08-899-575-154	Sequence 154, App
35	36.6	2.9	3282	1	US-08-899-575-154	Sequence 154, App
36	36.6	2.9	3282	5	PCT-US95-08743-154	Sequence 154, App
37	36.6	2.9	3282	5	PCT-US95-08743-154	Sequence 154, App
38	36.6	2.9	13254	1	US-08-276-852-156	Sequence 156, App
39	36.6	2.9	13254	1	US-08-276-852-170	Sequence 170, App
40	36.6	2.9	13254	1	US-08-899-575-156	Sequence 156, App
41	36.6	2.9	13254	1	US-08-899-575-156	Sequence 156, App
42	36.6	2.9	13254	1	US-08-899-575-156	Sequence 156, App
43	36.6	2.9	13254	1	US-08-899-575-170	Sequence 170, App
44	36.6	2.9	13254	5	PCT-US95-08743-156	Sequence 156, App
45	36.6	2.9	13254	5	PCT-US95-08743-170	Sequence 170, App

ALIGNMENTS

RESULT 1
US-09-348-930A-8
; Sequence 8, Application US/09348930A
; Patent No. 6573431

; GENERAL INFORMATION:

; APPLICANT: Lenee, P

; APPLICANT: Gruber, V

; APPLICANT: Baudino, S

; APPLICANT: Merot, B

; APPLICANT: Benicourt, C

; APPLICANT: Cudrey, C

; TITLE OF INVENTION: Recombinant Predodenal Lipases and Polypeptides Derivatives Prod

; FILE REFERENCE: 18433/2012

; CURRENT APPLICATION NUMBER: US/09/348,930A

; CURRENT FILING DATE: 1999-07-02

; PRIOR APPLICATION NUMBER: 08/945,321

; PRIOR FILING DATE: 1998-02-12

; PRIOR APPLICATION NUMBER: PCT/FR96/00606

; PRIOR FILING DATE: 1996-03-19

; NUMBER OF SEQ ID NOS: 37

; SOFTWARE: Patent in version 3.0

; SEQ ID NO 8

; LENGTH: 1367

; TYPE: DNA

; ORGANISM: Homo sapiens

; FEATURE:

; NAME/KEY: CDS

; LOCATION: (47)..(1240)

; US-09-348-930A-8

Query Match 37.0%; Score 470; DB 4; Length 1367;

Best Local Similarity 64.2%; Pred. No. 3.8e-142;

Mismatches 707; Conservative 0; Mismatches 395; Indels 0; Gaps 0;

QY	119	CTGTGACCCAGACGATTCATGATATATTAGTGAATATCATCAACATCAAGCTATCCCT	178
Db	123	CTGGAAGCCCTGAAGTGACTATGAACTATTAGTGAATATCATCAACATCAAGCTATCCCT	182
QY	179	GTGAGGAATATGAAGTCGCAACTGAAGATGGGTATATCTTCTGTAAACAGGATTCCTC	238
Db	183	ATGGAATATGAAGTGTGTGACTGTGAGATGGTATATCTTCTGTAAACAGGATTCCTC	242
QY	239	GAGCCCTAGTGCAACTAAGAGACAGAGTTCCAGGCTGTGGTGTGTACTGTGACGATGGCC	298
Db	243	ATGGGAAGAAAATTTCAGGGAATATACAGGACGACCTGTGTGTGTGTGTGTGTGTGT	302
QY	299	TAGTGTGAGGTGTAGCAACTGATTTCCAACTGCCCAACAAATAGCTGGGCTTCATTC	358
Db	303	TGCTTGATCAGCCCAAACTGATTTCCAACTGCCCAACAAATAGCTGGGCTTCATTC	362

359 TGGCAGATGCTGTTGTTGAGCTGTGGATGGGAAACAGCAGGGAAGAGCTGCTCGAA 418
 Db
 363 TGGCAGATGCTGTTGTTATGATGTGGCTGGGCAACAGCAGAGGAACACCTGGCCAGAA 422
 Qy 419 AACACAGACACTCTCCATAGACCAAGATGAGTCTGGGCTTTTCAGTTATGATGATGG 478
 Db 423 GAACTTGTACTATTACACAGATTCACTGAAATTCAGTGAATTCAGCTTTGATGAATGG 482
 Qy 479 CTAGTTTGAACCTTCCTGAGTGTATTAATTTTGCAGAAACGGCCAGGAAGA 538
 Db 483 CTAATATGACCTTCCAGCCACATCGACTTCATTTAAAGAAATCTGACAGACAGC 542
 Qy 539 TCTATTATGCGGTATTTCAAGGCAACCACTGAGCTTTTATGCAATTTTCACCATGC 598
 Db 543 TACACTATGTTGGCAATTCACGGCACCACCACTTGGTTTATTTGCTTTTCACCAATC 602
 Qy 599 CAGAGCTGGCTCAGAAATCAATATGATTTTCTGTTTACCCATCAGCAGCTTTAAGC 658
 Db 603 CCAGCTGGCTTAAAGAAATCAAACTTCTATGCTCTAGCTCTGTTGGCAGCTGAGT 662
 Qy 659 ATGCAAAAGCCCGGAGCAAAATTTTGTGCTGCCAGATATGATGATCAAGGATTTG 718
 Db 663 ATACAAAGGCTTATTAAGCAAACTTAGATTTGTTCTCAATCCCTCTTCAAGTTATAT 722
 Qy 719 TTGCAAAAGAAATTTCTGTATCAGACAGATTTCTCAGCAAACTTGTATTTACCTTT 778
 Db 723 TTGCTGACAAATATTTACCCACAACTTCTTGTGATCAATTTCTGCTACTGAAAGTGT 782
 Qy 779 GTGCGCAGGTGATTTGATCAGATTTGATGATATCATGTTACTTCTGGTGGATTC 838
 Db 783 GCTCCCTGAGATGCTGNAATCTCTTTGAGCAATGCTTATTTAATTTGATTTG 842
 Qy 839 ACACCAACATATGAACATGAGCGGAGCAAGTGTATGCTGCCACACACTTTCTGCTG 898
 Db 843 ACAGTAAGAACTTTAAACAGAGTCTGTTGGATGTGTATCTATCACATAATCCAGCAG 902
 Qy 899 CATCTGTGCAAAATTTCTACCTGAGCGAGCAGGAGTATCTGTTGTAAGTCTCGG 958
 Db 903 CTTCCTGTTCAAAACATGTTCCATTTGACCCAGGCTGTTAAGTCTGGAAATTCAGCT 962
 Qy 959 TTGACTGGGAGTGAGACCAAAATCTGGAATAATGCAATCAAGCCACTCTCTGTAAGT 1018
 Db 963 ATGACTGGGAAGCCAGTTGAGAAATAGGATGCACTATGATGATGCTCCACCTCCCT 1022
 Qy 1019 ACAGATCAGAGATGACGCTCCCTACAGCAATGTGGCAGGAGTCAAGCTGCTTT 1078
 Db 1023 ACAATGTGACGCCATGAATGTACCAATTTGAGTGTGGAACGTTGGCAAGGACTGTTG 1082
 Qy 1079 CAAATCCAGAGAGCTGGAATAATGCTGCTCTGAGGTGACCAACCTCATCTACCAATA 1138
 Db 1083 CTGACCCCAAGATGTTGGGCTTTTGTCTTCCAAACTCCCAATCTTATTTACCAAG 1142
 Qy 1139 ATATTCCTGATGGCTCAGTGATTTCTGCTGGGTTTGGATGCTCTCCTCAGCTATGT 1198
 Db 1143 AGATTCCTTTTACAATCACTTGACCTTATCTGCGCAATGGAATGCGCCCTCAAGAGT 1202
 Qy 1199 ACAATGAATCATCCATCTGAT 1220
 Db 1203 ACAATGACATGTTTCTATGAT 1224

RESULT 2
 US-09-186-489-3
 ; Sequence 3, Application US/09186489
 ; Patent No. 6375947
 ; GENERAL INFORMATION:
 ; APPLICANT: Bolen, Paul L.
 ; APPLICANT: Scharf Jr., Lewis G.
 ; TITLE OF INVENTION: Purified Recombinant Kid Pregastric Esterase, and
 ; TITLE OF INVENTION: Processes for its Production and Use
 ; FILE REFERENCE: 5499/3

; CURRENT APPLICATION NUMBER: US/09/186,489
 ; CURRENT FILING DATE: 1998-11-05
 ; NUMBER OF SEQ ID NOS: 7
 ; SOFTWARE: PatentIn Ver. 2.0
 ; SEQ ID NO 3
 ; LENGTH: 1411
 ; TYPE: DNA
 ; ORGANISM: Kid (Goat)
 ; US-09-186-489-3

Query Match 36.0%; Score 456.8; DB 4; Length 1411;
 Best Local Similarity 62.0%; Pred. No. 7.4e-138; Indels 3; Gaps 1;
 Matches 741; Conservative 0; Mismatches 452;
 Qy 42 AATGAAATGTGGCTTCTGATTTCTGGTGGGCTATATTTCCAGAGAAATGTAATTCAGT 101
 Db 61 ATTGAAATGTGGTGGCTTCTGATTTCTGTAACGGTGTGTTTCCATCCACATGCTCGAAATGCATT 120
 Qy 102 ACATATGC---CACTAAAGCTGTGGACCCAGAGCAATTCATGATATTTAGTGAATCAT 158
 Db 121 TTGTTTCTTGGAAAAATTTGCTAAGAACCTTGAAGCCAGTATGATGTGATGCATGAT 180
 Qy 159 CCAACATCAAGGCTATCCCTGTGAGGAATATGAAGTCGCAACTGAAGATGGGTATATCTCT 218
 Db 181 TTCTTCTTGGGGCTACCCAGTGAATGCAATAAGTTTAACTGCAGATGCTATATCTCT 240
 Qy 219 TTCTGTTAAAGATTTCTCGAGGCTTAGTGCAACCTTAAGAGACAGGTTCCAGGCTGT 278
 Db 241 TCAGGTCTATCGGATTTCTCATGGAAGAATGATGCTTAATCATTTAGGTGAGAGACCTGT 300
 Qy 279 GGTGTTCTGAGCATGCTTGTGAGGCTGTGAGCACTCTGATTTCCAACTGCCCAA 338
 Db 301 TGTGTTCTGAGCAATGCTTCTTGTGCTCAGCTACAACTGATTTCCAACTTCCCAA 360
 Qy 339 CAATAGCTGGCTTCAATTTGCGAGATGCTGTTTTGACGTGTGAGTGGGAAACAGCAG 398
 Db 361 CAACAGCTGGCTTCTCTCTGCGAGATGCTGTTTATGAGCTGTGCTGGGAAACAGCAG 420
 Qy 399 GGGAAAGCTTCTCGAAACACAGACACTCTCCATAGACCAAGATGATGTTCTGGGC 458
 Db 421 AGGAACACTTGGGCCCCAGGAACATTTATCTATTCACAGACTCCCTCGAATTTCTGGGC 480
 Qy 459 TTTCAAGTTATGATGAGATGCTAGGTTTGAAGCTTCTGCGAGTGATAAACTTTATTTGCA 518
 Db 481 TTTCAAGCTTTGATGAATGCTGAAATATGATGCTTCCATCTACAATGATTTCTTTAAA 540
 Qy 519 GAAACGGGCCAGGAAGATCTATTTATGCTGGCTATTTCAAGGGCACCACCATGGGCTT 578
 Db 541 GAGAACAGGACAGGAAGAGTACACTATGTTGGCCATTTCCCAAGGCCACCATTTGGTTT 600
 Qy 579 TATTGCTATTTCCACCATGCGCAGAGCTGGCTCAGAAAAATCAAAATGATTTTGTCTTTCAG 638
 Db 601 TGTGCGCTTTTCTACCAATCCCACTGGCTGAAATTCGAAATGCTTCCATGCAATGATC 660
 Qy 639 ACCCAATAGCTCTGTTAAGCATGCAAAAGCCCGGAGCAAAATTTTGTGCTGCGAGA 698
 Db 661 CCCAGTCCCAAGTGAAGCACCAGAGCTGTTTAAACAACTTGCACATTTATTCCTCA 720
 Qy 699 TATGATGATCAAGGATTTGTTGGCAAAAGATTTCTGTATCCAGACAGATTTCTCAG 758
 Db 721 CTTCTCTTCAAGATTTATTTGTTAACAATGTTCTACCCACAAATTTTTTTGACA 780
 Qy 759 ACAACTTGTATTTTACCTTTTGTGCGCAGGTGATTTCTGATCAGATTTGTAGTAAATCAT 818
 Db 781 ATTTCTTGGTGTGAAGTGTCTCTCGTGAAGACACTGATGCTCTTTGTAAGAATGCTT 840
 Qy 819 GTTACTTCTGGTGGATTCACACCAATATGAATGATGAGCCGAGCAAGTGTATATGC 878
 Db 841 GTTTCGCAATTTACTGAGCTGCAATAAATACTTCAATGATGCTGCTGTAGATGTATGT 900
 Qy 879 TGCCACACTCTTGTGGAACATCTGTGCAAAATATTTCTACACTGGAGCCAGGAGTAA 938
 Db 901 AGCATAATCCAGCAGGAGCTTCTGTTCAAAACATCTCTCCACTGGAGACAGGCTATTA 960

QY 939 TTCTGGTGAACCTCCGGGCAATTTGACTGGGGAGTGCAGACCAAAATCTGGAATAATGCCAA 998
 Db 961 GTCTGGGAAATTCGAAGCTTTTGACTGGGGAGCCTCACTTGAGAACCTTAATGCAATATAA 1020
 QY 999 TCAGCAACCTCTGTAAGGTACAGAGTGCAGAGATAGCGGTCCCTCAGCAAAATGTGGAC 1058
 Db 1021 TCAGCCCAACCTCCCATCTCAATTAATTAACAGCCATGAATGTCCTCAATGAGATGGAG 1080
 QY 1059 AGGAGTTCAGGACTGGCTTCAATCCAGAGACGTGAAATGCTGCTCTCTGAGGTGAC 1118
 Db 1081 TGCTGGCCCAAGACCTGTTGGCTGACCCCTCAGGATGTTGACCTTTTGGCTTTCAAAACTCTC 1140
 QY 1119 CAACCTCATCTACCAATAAGAAATATTCCTGAATGGGCTCAGGTGGAATTTCACTGGGGTTT 1178
 Db 1141 TAATCTCATTCACCAAGAAATTCCTGAATTTACATCATCTGGACTTTTCTGCGCAAT 1200
 QY 1179 GAATGCTCTCCAGGTATGTAACAATGAATCATCCATCTGATGACAGAGGAGGA 1234
 Db 1201 GGATGACCTCAAGAAGTTTACAATGAATTTATTTCTTTGATGGCAAAAGACAAA 1256

RESULT 3

US-10-043-665B-3

; Sequence 3, Application US/10043665B

; Patent No. 6582948

; GENERAL INFORMATION:

; APPLICANT: Bolen, Paul L.

; APPLICANT: Cihak, Paul L.

; APPLICANT: Scharpf Jr., Lewis G.

; TITLE OF INVENTION: Recombinant Kid Pregastric Esterase and Methods for Its

; FILE REFERENCE: IFF-0009

; CURRENT APPLICATION NUMBER: US/10/043,665B

; PRIOR FILING DATE: 2002-09-24

; PRIOR FILING DATE: 1998-11-05

; NUMBER OF SEQ ID NOS: 7

; SOFTWARE: Patent in Ver. 2.0

; SEQ ID NO 3

; TYPE: DNA

; ORGANISM: Kid (Goat)

US-10-043-665B-3

Query Match

Best Local Similarity 36.0%; Score 456.8; DB 4; Length 1411;

Matches 741; Conservative 0; Mismatches 452; Indels 3; Gaps 1;

QY 42 AATGGAATGGCTTCTGATTTCTGGTGGCGTATATGTTCCAGAGAAATGTAATTCAGT 101
 Db 61 ATTGAAATGGTGGCGTACTGTAACTGGTGTGTTTCACTCCATGTTCTGGAATGCCAT 120
 QY 102 ACATATGC---CAACTAAAGCTGTGGACCCAGAGCAATTCATGAATATTAAGTGAATCAT 158
 Db 121 TTGTTTCTTGGAAATTTGCTAAGAACCTTGAGCCAGTATGAATGTCAGTCAATGAT 180
 QY 159 CCAACATCAAGGTATCCCTGTGAGGAATATGAAGTCGCAACTGAGATGGGTATATCCT 218
 Db 181 TTCTTTCTGGGGTACCCCAAGTGAATGATGATTAAGTTAATCACTGAGATGGGTATATCCT 240
 QY 219 TTCTGTTAAAGGATTCCTCGAGGCTAGTGAACCTAAGAGACAGAGTTCAGGCGCTGT 278
 Db 241 TCAGGTCTATCGAATTCCTCATGGAAGAAATGATGCTATCATTTAGTTCAGAGACCTGT 300
 QY 279 GGTGTTTACTGCAGTATGGCTAGTTGAGGTGCTAGCAACTGGAATTTCAACCTGCCAA 338
 Db 301 TGTGTTTCTGCAGTATGCTCTTCTGCTCAGCTACAAACTGGAATTTCCAACTTTCCAA 360
 QY 339 CAATAGCTGGGCTTCATTTCTGCAGATGCTGGTTTGAAGTGTGATGGGGAACAGCAG 398
 Db 361 CAACAGCTGGGCTTCTCTCTGTCAGATGCTGGTTATGACGTGTGGCTGGGGAACAGCAG 420

RESULT 4

US-09-186-489-4

; Sequence 4, Application US/09186489

; Patent No. 6375947

; GENERAL INFORMATION:

; APPLICANT: Bolen, Paul L.

; APPLICANT: Cihak, Paul L.

; APPLICANT: Scharpf Jr., Lewis G.

; TITLE OF INVENTION: Purified Recombinant Kid Pregastric Esterase, and

; TITLE OF INVENTION: Processes for its Production and Use

; FILE REFERENCE: 5499/3

; CURRENT APPLICATION NUMBER: US/09/186,489

; CURRENT FILING DATE: 1998-11-05

; NUMBER OF SEQ ID NOS: 7

; SOFTWARE: Patent in Ver. 2.0

; SEQ ID NO 4

QY 399 GCGAAGCGCTGGTCTCGAAACACAGACACTCTCCATAGACCAAGATGAGTCTTGGGC 458
 Db 421 AGGAACACTTGGGCCCAGAACATTTATCTATTCCAGACTCCCTGGAATCTGGGC 480
 QY 459 TTTTCAGTTATGATGAGATGGCTAGTGTGACCTTCTGCAAGTATAAACTTTATTTGCA 518
 Db 481 TTTTCAGTTATGATGAAATGGCTGAATATGACCTTCCATCTACAATTTGATTTCTATAA 540
 QY 519 GAAACGGGCCAGGAAAGATCTATTTATGTCGCTATTTCACAGGACCAACCATGGCTT 578
 Db 541 GAAACAGGACAGAAAGCTACATATGTTGGCCATCCCAAGGACCAACCATGGCTT 600
 QY 579 TATTGATTTTCCACCATGCCAGCTGGCTCAGAAATCAAAATCTATTTTGTCTTTAGC 638
 Db 601 TGTGCGCTTTTCTACCAATCCACACTGGCTGAAATAATCGAAGTCTTCCATGCAATGAC 660
 QY 639 ACCATAGCACTGTTAAGCATGCAAAAGCCCGGACCAAAATTTTGTGCTGCGAGA 698
 Db 661 CCAGTGGCCACAGTGAAGCACCAGAGCCTGTTTAACTTTGCACTTTATTTCTCA 720
 QY 699 TATGATGATCAAGGGATTTGTTGGCAAAAGAAATTTCTGTATCAGACAGCAATTTCTCAG 758
 Db 721 CTTCCTCTTCAAGATTAATTTGTTGTAACAAATGTTCTACCCACACAATTTTGTGAACA 780
 QY 759 ACAACTGTTTATTTCCTTTGTCGCGAGGTGATCTTGTATCAGATTTGTAGTAATCAT 818
 Db 781 ATTTCTTGGTGTGAAGTGTGCTCTCTGAGACACTGATGCTCTTTGTAAGAAATGCTTT 840
 QY 819 GTTACTTCTGGTGGATTCACACCAACCAATATGAACTGAGCCGAGCAAGTGTATATGC 878
 Db 841 GTTTGCCATTTACTGGAGTGCACAAATAAACTTCAACATGAGTGGCTTGAATGTATGT 900
 QY 879 TGCCCAACACTCTTGTGCGAATCTCTGCAAAATATTTTACACTGAGCCAGCAGTGA 938
 Db 901 AGCACATAATCCAGCAGGAGCTTCTGTTCAAAACATCTCCCACTGGAGACAGGCTATTAA 960
 QY 939 TTCTGTTGAATCTCGGGCACTTTCAGCTGGGGAGTGAACCAAAATCTGGAATAATGCA 998
 Db 961 GTCTGGGAAATTCGAAGCTTTTGTGCTGGGGAGGCTCAGTTGAGAACCTTAATGCAATATA 1020
 QY 999 TCAGCCAACTCTGTAGGTACAGAGTACAGAGATATGAGGTCCTCTACAGCAATGTGGAC 1058
 Db 1021 TCAGCCCAACCTCTCCCATCTACAAATTTACAGCCATGAATGTCCTCAATTTGAGTGGAG 1080
 QY 1059 AGGAGTTCAGACTGGCTTTCAATTCAGAAAGCTGGAATAATGCTCTCTGTAGGTGAC 1118
 Db 1081 TGCTGGCCAAAGACTGTTGGCTCAGCTCAGGATGTTGACCTTTTGTCTTTCAAAACTCTC 1140
 QY 1119 CAACCTCATCTACCAATAAGAAATTTCTGAAATGGGCTCAGGTGATTTTCATCTGGGTTT 1178
 Db 1141 TAATCTCATTCACCAAGGAAATTCGAATTAATCATCTGAGACTTTATCTGGCAAT 1200
 QY 1179 GGATGCTCTCAGCGTATGTACAAATGAAATCATCCATCTGATGACAGAGGAGGA 1234
 Db 1201 GGATGCACTCAAGAGTTCACAAATGAAATTTATTTCTTGTGCAAAAGACAAA 1256

LENGTH: 1134
TYPE: DNA
ORGANISM: Bovine
US-09-186-489-4

Query Match 35.7%; Score 453.6; DB 4; Length 1134;
Best Local Similarity 62.7%; Pred. No. 7.1e-137;
Matches 705; Conservative 0; Mismatches 419; Indels 0; Gaps 0;

QY 111 AACTAAGCTGTGGACCCAGAACATTCATGATATATAGTGAATCATCAACATCAAGG 170
DB 9 AAAATTTGCTAAGAACCCCTGAAGCCAGTATGAATTTAGTCAGATGATTTCTACTGGGG 68

QY 171 CTATCCCTGTGAGGAATATGAAGTCGCAACTGAAGATGGGTATATCTTTCTGTAAACAG 230
DB 69 CTACCAAGTGAGATGATGATTAAGTTATACTGGGATGGTTATCTTCAGGTCTATCG 128

QY 231 GATTCCTCGAGGCTAGTGCACCTTAAGAAGACAGGTTCCAGGCTGTGGTTACTGCA 290
DB 129 GATTCCTCGAGGCTAGTGCACCTTAAGAAGACAGGTTCCAGGCTGTGGTTACTGCA 188

QY 291 GCATGGCTAGTTGGAGGTCTAGCAACTGGAATTTCCAACTCCCAACAATAGCCTGGG 350
DB 189 GCATGGCTAGTTGGAGGTCTAGCAACTGGAATTTCCAACTCCCAACAAGCCTGGG 248

QY 351 CTTCAATTCGTGAGATGCTGTTTGGATGGGATGGGAAACAGGCGCTG 410
DB 249 CTTCAATTCGTGAGATGCTGTTTGGATGGGATGGGAAACAGGCGCTG 308

QY 411 GTCTCGAATAACACAGACACTCTCCATAGACCAAGATGAGTCTGTGGCTTTCAGTTATGA 470
DB 309 GGCCAGGAACTTATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTA 368

QY 471 TGAGATGGCTAGTTGACCTCTCGAGTGAATAAATTTATTTTGGAGAAACGGGCCA 530
DB 369 TGAATGGCGGAAATGATGACCTTCTCACTAATTAATTAATTAATTAATTAATTAATTA 428

QY 531 GGAAAGATCTATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTA 590
DB 429 GAAGAAGCTTACATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTA 488

QY 591 CACCATGCCAGAGCTGGCTCAGAAATCAGAAATCAGAAATCAGAAATCAGAAATCAGAA 650
DB 9 AAAATTTGCTAAGAACCCCTGAAGCCAGTATGAATTTAGTCAGATGATTTCTACTGGGG 68

QY 171 CTATCCCTGTGAGGAATATGAAGTCGCAACTGAAGATGGGTATATCTTTCTGTAAACAG 230
DB 69 CTACCAAGTGAGATGATGATTAAGTTATACTGGGATGGTTATCTTCAGGTCTATCG 128

QY 231 GATTCCTCGAGGCTAGTGCACCTTAAGAAGACAGGTTCCAGGCTGTGGTTACTGCA 290
DB 129 GATTCCTCGAGGCTAGTGCACCTTAAGAAGACAGGTTCCAGGCTGTGGTTACTGCA 188

QY 291 GCATGGCTAGTTGGAGGTCTAGCAACTGGAATTTCCAACTCCCAACAATAGCCTGGG 350
DB 189 GCATGGCTAGTTGGAGGTCTAGCAACTGGAATTTCCAACTCCCAACAAGCCTGGG 248

QY 351 CTTCAATTCGTGAGATGCTGTTTGGATGGGATGGGAAACAGGCGCTG 410
DB 249 CTTCAATTCGTGAGATGCTGTTTGGATGGGATGGGAAACAGGCGCTG 308

QY 411 GTCTCGAATAACACAGACACTCTCCATAGACCAAGATGAGTCTGTGGCTTTCAGTTATGA 470
DB 309 GGCCAGGAACTTATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTA 368

QY 471 TGAGATGGCTAGTTGACCTCTCGAGTGAATAAATTTATTTTGGAGAAACGGGCCA 530
DB 369 TGAATGGCGGAAATGATGACCTTCTCACTAATTAATTAATTAATTAATTAATTAATTA 428

QY 531 GGAAAGATCTATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTA 590
DB 429 GAAGAAGCTTACATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTA 488

QY 591 CACCATGCCAGAGCTGGCTCAGAAATCAGAAATCAGAAATCAGAAATCAGAAATCAGAA 650
DB 9 AAAATTTGCTAAGAACCCCTGAAGCCAGTATGAATTTAGTCAGATGATTTCTACTGGGG 68

QY 1071 CTGGCTTTCAAAATCCAGAGACGTGAAATGCTGCTCTCTGAGGTGACCAACCTCATCTA 1130
DB 969 CTTGTTGGCTGACCCCTCAGGATGTTGACTTTCTGCTTTCAAAACTCTCTAATCTCA 1028

QY 1131 CCATAAGATATTTCTGAAATGGGCTCAGTGGATTTCACTGCGGTTTGGATGCTCTCA 1190
DB 1029 CCACAAGGAAATTCAAAATTAACAATCACTTTGGACTTTATCTGGCAATGGATGACCTCA 1088

QY 1191 CCGTATGTATCAATCAAAATCATCTCATCTGATGACGAGGAGGAGA 1234
DB 1089 AGAAGTTTACATGAAATTTGTTCTTGTATGGCCGAGACAAA 1132

RESULT 5

US-10-043-665B-4
; Sequence 4, Application US/10043665B
; Patent No. 6582948
; GENERAL INFORMATION:
; APPLICANT: Boehr, Paul L.
; APPLICANT: Cihak, Paul, L.
; TITLE OF INVENTION: Recombinant Kid Prestaeric Esterase and Methods for Its
; FILE REFERENCE: IPF-0009
; CURRENT APPLICATION NUMBER: US/10/043,665B
; PRIOR FILING DATE: 2002-09-24
; PRIOR APPLICATION NUMBER: US 09/186,489
; NUMBER OF SEQ ID NOS: 7
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 4
; LENGTH: 1134
; TYPE: DNA
; ORGANISM: bovine
US-10-043-665B-4

Query Match 35.7%; Score 453.6; DB 4; Length 1134;
Best Local Similarity 62.7%; Pred. No. 7.1e-137;
Matches 705; Conservative 0; Mismatches 419; Indels 0; Gaps 0;

QY 111 AACTAAGCTGTGAGCCAGAACATTCATGATATATAGTGAATCATCAACATCAAGG 170
DB 9 AAAATTTGCTAAGAACCCCTGAAGCCAGTATGAATTTAGTCAGATGATTTCTACTGGGG 68

QY 171 CTATCCCTGTGAGGAATATGAAGTCGCAACTGAAGATGGGTATATCTTTCTGTAAACAG 230
DB 69 CTACCAAGTGAGATGATGATTAAGTTATACTGGGATGGTTATCTTCAGGTCTATCG 128

QY 231 GATTCCTCGAGGCTAGTGCACCTTAAGAAGACAGGTTCCAGGCTGTGGTTACTGCA 290
DB 129 GATTCCTCGAGGCTAGTGCACCTTAAGAAGACAGGTTCCAGGCTGTGGTTACTGCA 188

QY 291 GCATGGCTAGTTGGAGGTCTAGCAACTGGAATTTCCAACTCCCAACAATAGCCTGGG 350
DB 189 GCATGGCTAGTTGGAGGTCTAGCAACTGGAATTTCCAACTCCCAACAAGCCTGGG 248

QY 351 CTTCAATTCGTGAGATGCTGTTTGGATGGGATGGGAAACAGGCGCTG 410
DB 249 CTTCAATTCGTGAGATGCTGTTTGGATGGGATGGGAAACAGGCGCTG 308

QY 411 GTCTCGAATAACACAGACACTCTCCATAGACCAAGATGAGTCTGTGGCTTTCAGTTATGA 470
DB 309 GGCCAGGAACTTATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTA 368

QY 471 TGAGATGGCTAGTTGACCTCTCGAGTGAATAAATTTATTTTGGAGAAACGGGCCA 530
DB 369 TGAATGGCGGAAATGATGACCTTCTCACTAATTAATTAATTAATTAATTAATTAATTA 428

QY 531 GGAAAGATCTATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTA 590
DB 429 GAAGAAGCTTACATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTA 488

QY 591 CACCATGCCAGAGCTGGCTCAGAAATCAGAAATCAGAAATCAGAAATCAGAAATCAGAA 650
DB 9 AAAATTTGCTAAGAACCCCTGAAGCCAGTATGAATTTAGTCAGATGATTTCTACTGGGG 68

Db 489 TACCAAGTCCACATTTGGTGTAAATCAAGTCTTCTATGATAGCCAGTTGGCAC 548
 Qy 651 TGTAAAGCATGCAAAAGCCGGGACCAAAATTTTGTGTGCGCAGATATGATGACAA 710
 Db 549 AGTGAAGTACACCAAGAGCGCTGTTTAAACAACTTGCACCTTATCTCCTCTTCAA 608
 Qy 711 GGGATTGTTGGCAAAAGAAATTTCTATACAGACCAAGATTTCTCAGACAACTTGTAT 770
 Db 609 GATTATATTTGGTGCAAAATGTTCTACCCACACACTTTTGGAAACAAATTTCTGGT 668
 Qy 771 TTACCTTTTGGCCAGGTGATTTCTTGTATGATGATTTGTATGATATATCATTTCTGGG 830
 Db 669 TGAATATGCTCCCGTGAGACACTGATGTCCTTTGTAAAGATGCTTGTGTCATTAC 728
 Qy 831 TGGATTCAACACCAACATATGAACATGAGCGGAGCAGTGTATGTCGCCACACTCT 890
 Db 729 TGGATTGACATAAACTTCAACATGAGTGTCTTGTATGATGATATGATGATATATCC 788
 Qy 891 TGTGGAACATCTGTGCAAAATTTCTATACAGTGGAGCCAGGAGTGAATTTCTGGTGAAT 950
 Db 789 AGCAGGAACTTCTGTTCAAAACACCCCTCCACTGGAGACAGGCTGTTAAGTCTGGAAAT 848
 Qy 951 CCGGCATTTGACTGGGAGTGAGACCAAAATCTGGAATAATGCAATGCAATGCAATGCA 1010
 Db 849 CCAAGCTTTGACTGGGAGGCGCCATATCAGAACCTTAATGCAATATCATCAGCCACAC 908
 Qy 1011 TGTAAAGTACAGAGTACAGATATGACGCTCCCTACAGCAATGTGGAGGAGTCCAGGA 1070
 Db 909 TCCCATCTCAATTTAAACGCCATGATGTCCTCAATGTCAGTGTGGAGTGTGACATGA 968
 Qy 1071 CTGGCTTTCAATCCAGAGAGTGAATAATGCTCTCTGAGGAGTGAACCACTCATCTA 1130
 Db 969 CCGTGTGGCTGACCTCAGGAGTGTGACTTTCTGCTTTCAAAACTCTCTAATCTCATTTA 1028
 Qy 1131 CCATAGGAATATTTCTGATGGCTCAGTGTGATTTCTGAGTGTGAGTGTGAGTGTGAGT 1190
 Db 1029 CCACAGGAATTTCAAAATTAATCAATCTTGGACTTTTATCTGGGCAATGGATGACCTCA 1088
 Qy 1191 CGGTATGTACATGAATCAATCATCCATCTGATGAGGAGGAGGA 1234
 Db 1089 AGAAGTTTACATGAATTTGTTCTTGTATGTCGCCAGACACAAA 1132

RESULT 6

US-09-186-489-1
 ; Sequence 1, Application US/09186489
 ; Patent No. 6375947
 ; GENERAL INFORMATION:
 ; APPLICANT: Bolen, Paul L.
 ; APPLICANT: Chak, Paul L.
 ; APPLICANT: Scharpf Jr., Lewis G.
 ; TITLE OF INVENTION: Purified Recombinant Kid Pregastric Esterase, and
 ; TITLE OF INVENTION: Processes for its Production and Use
 ; FILE REFERENCE: 5499/3
 ; CURRENT APPLICATION NUMBER: US/09/186,489
 ; CURRENT FILING DATE: 1998-11-05
 ; NUMBER OF SEQ ID NOS: 7
 ; SOFTWARE: Patentin Ver. 2.0
 ; SEQ ID NO 1
 ; LENGTH: 1134
 ; TYPE: DNA
 ; ORGANISM: Kid (Goat)
 ; FEATURE:
 ; NAME/KEY: CDS
 ; LOCATION: (1)..(1134)
 ; US-09-186-489-1

Query Match 35.5%; Score 450.4; DB 4; Length 1134;
 Best Local Similarity 62.5%; Pred. No. 7,7e-136;
 Matches 703; Conservative 0; Mismatches 421; Indels 0; Gaps 0;
 Qy 111 AACTAAAGCTGTGGACCCAGAGCAATTCATGAATATTAGTGAATCATCCACATCAAGG 170

Db 9 AAAAAATTGCTAAGAACCTTGAAGCCAGTATGAATGTGAGTCAGATGATTTCTCTCTGGG 68
 Qy 171 CTATCCCTGTGAGGAATATGAAGTGCAGAACTGAAAGTGGTATATCTTTCTGTGTAAACAG 230
 Db 69 CTACCCAAAGTGAAGTGCATAAAGTTATTAAGTGCAGATGGCTATATCTTTCAGGTCTATCG 128
 Qy 231 GATTCCTCTGAGGCTAGTGCACAACTTAAGAGACAGGTTCACAGGCTGTGGTGTCTACTGCA 290
 Db 129 GATTCCTCATGGAAGAATGATGCTAATCATTTAGGTAGAGACCTGTGTGTGTCTGCA 188
 Qy 291 GCATGCCCTAGTTGGAGGTGCTAGCAACTGGAATTTCCAACTGCCCCCAACATAGCCTGGG 350
 Db 189 GCATGGCTCTTCTGCTCCAGCTACAACTGGAATTTCCAACTTCCCACTTCCCAACACAGGCTGGG 248
 Qy 351 CTTTCATCTTGGCAGAGTGTGTTTGAAGTGTGGATGGGGAACAGCAGGGGAAACGCCCTG 410
 Db 249 CTTTCCTCTGGCAGATGCTGTTTATGACGTGTGGTGGGGAACAGCAGGGGAAACACTTG 308
 Qy 411 GTCTCGAAACCAAGACACCTCTCCATAGCAAGATGATTTCTGGGCTTTTCAGTTATGA 470
 Db 309 GGGCCAGGAACATTTTATCTATTCCAGACCTCCCTGAAATTCGGGCTTTTCAGCTTTGA 368
 Qy 471 TGAGATGGCTAGGTTTGAACCTTCTGTCAGTGAATACTTTATTTTGCAGAAACCGGCCA 530
 Db 369 TGAATGGGCTGAATATGACCTTCCATCTACAATGATTTTCATCTTAAAGAGAACAGGACA 428
 Qy 531 GGAAGAATCTATTATGTGGCTATTTCACAGGGCCACCACTGGGCTTTTATTTGCAATTTTC 590
 Db 429 GAAGAAGTACTACTATGTTGGCCATTTCCCAAGGCACCACTTGGTTTGTGCGCTTTTC 488
 Qy 591 CACCATGCGAGAGTGGCTCAGAAAATGATTTTGTGCTTTAGCAGCCCATAGCCAC 650
 Db 489 TACCAATCCCACTGGCTGAAAAAATCGAAGTCTTCCATGCAATGAGCCCACTGCGCCAC 548
 Qy 651 TGTAAAGCATGCAAAAGCCCGGAGCAAAATTTTGTGCTGCCAGATATGATGATCAA 710
 Db 549 AGTGAAGCACACCCAGAGCTGTTTAAACAACTTGCACTTATCTCTCTCTCTCTCAA 608
 Qy 711 GGGATTGTTTGGCAAAAGAAATTTCTGATCAGACCAAGATTTCTCAGACAACTTGTAT 770
 Db 609 GATTATATTGGTAAACAAATGTTTACCCACACAAATTTTGTGAACAAATTTCTTGTGT 668
 Qy 771 TTACCTTTTGGCCAGGTGATTTCTGATCAGATTTGTAGTATATATCATGTTACTTCTGGG 830
 Db 669 TGAAGTGTCTCTCTGAGACACTGATGCTCTTTTGAAGATGCTTTTGTGCAATTAC 728
 Qy 831 TGGATTCAACACCAACATATGAACATGAGCCAGCAAGTGTATATGCTGCCCACTCT 890
 Db 729 TGGAGCTGACATAAAACTTCAACATGAGTGTCTGATGATGATGATGATGATGATGATGAT 788
 Qy 891 TGTGGAACATCTGTGCAAAATTTCTACACTGAGGCGAGGAGTGAATTTCTGTGGAAT 950
 Db 789 AGCAGGAGCTTCTGTTCAAAACATCTCCACTGGAGACAGGCTATTTAAGTCTTGGGAAAT 848
 Qy 951 CCGGCATTTGACTGGGAGTGAAGACCAAAATCTGGAATAATGCAATGCAATGCAATGCA 1010
 Db 849 CCAAGCTTTGACTGGGAGGCTCAGTGTGAACCTTAATGCAATTAATCAACCCACAC 908
 Qy 1011 TGTAAAGTACAGAGTACAGATATGAGGCTCCCTACAGCAATGTGGAGGAGTGAAG 1070
 Db 909 TCCCATCTCAATTTAAACGCCATGATGTCCTCAATTTGCAATGAGTGTGCGGCAAGA 968
 Qy 1071 CTGGCTTTCAATCCAGAGAGTGAATAATGCTCTCTGAGGAGTGAACCACTCATCTA 1130
 Db 969 CTTGTTGGCTGACCTCAGGATGTTGACCTTTTGTCTTTCAAACTCTCTAATCTCATTTA 1028
 Qy 1131 CCATAAAGATATTTCTGAAATGGGCTCACGTTGGAATTTTCATCTGGGCTTTTGGATGCTCTCA 1190
 Db 1029 CCACAGGAATTTCCAAATTAATCAATCATCTGAGCTTTTATCTGGCAATGATGATGATGAT 1088
 Qy 1191 CGGTATGTACATGAATCAATCATCCATCTGATGAGGAGGAGGA 1234

Db 1089 AGAAGTTTCAATGAAATTTCTTTGATGGCAAGACAAAA 1132

RESULT 7

US-10-043-665B-1
; Sequence 1, Application US/10043665B
; Patent No. 6582948

; GENERAL INFORMATION:
; APPLICANT: Bolen, Paul L.
; APPLICANT: Cihak, Paul, L.

; TITLE OF INVENTION: Recombinant Kid Pregastric Esterase and Methods for Its
; TITLE OF INVENTION: Production and Use

; FILE REFERENCES: IFF-0009

; CURRENT APPLICATION NUMBER: US/10/043,665B
; CURRENT FILING DATE: 2002-09-24

; PRIOR APPLICATION NUMBER: US 09/186,489
; PRIOR FILING DATE: 1998-11-05

; NUMBER OF SEQ ID NOS: 7
; SOFTWARE: PatentIn Ver. 2.0

; SEQ ID NO 1

; LENGTH: 1134

; TYPE: DNA

; ORGANISM: Kid (Goat)

; FEATURE:

; NAME/KEY: CDS

; LOCATION: (1)..(1134)

US-10-043-665B-1

Query Match 35.5%; Score 450.4; DB 4; Length 1134;

Best Local Similarity 62.5%; Pred. No. 7.7e-136; Indels 0; Gaps 0;

Matches 703; Conservative 0; Mismatches 421; Indels 0; Gaps 0;

Qy 111 AACTAAAGCTGTGGACCCAGAGCAATTCATGAATATTAGTGAATCATCAACATCAAGG 170

Db 9 AAAAATTTGCTAAGAACCTCGAAGCCAGTATGAATGTGAGTCAGATGATTTCTTCCTGGGG 68

Qy 171 CTATCCCTGTGAGGAATATGAGTGCAGCTGAGATGGGTATATCTCTTCTGTTAAACAG 230

Db 69 CTACCCAGTGTAGATGATGAAGTTAATCTGAGATGGGTATATCTCTCAGGTCTATCG 128

Qy 231 GATTCCTCGAGGCTTAGTGCACCTAAGAAAGACAGGTTCCAGGCTCTGTGTTACTGCA 290

Db 129 GATTCCTCATGGAAGATGATGCTAATCATTTAGGTTCAGAGACCTGTTGTTCTGCA 188

Qy 291 GCATGGCTAGTTGGAGGTGTAGCAACTGATTTCCAACTGCTGCCCAACATAGCTGGG 350

Db 189 GCATGGCTCTTCTTCTCCCTCAGCTACAAACTGGATTTCCAACTTCCCAACACAGCCTGGG 248

Qy 351 CTTCATTTCTGCAGATGCTGTTTGTGACGTGTGGATGGGAAACAGAGGGGAAACGCTTG 410

Db 249 CTTCCTCTGCAGATGCTGTTTGTGACGTGTGGATGGGAAACAGAGGGGAAACGCTTG 308

Qy 411 GTCTCGAAACACAGACATCTCTCCATAGACCAAGATGATTTCTGGGCTTTCAGTTATGA 470

Db 309 GGCCCGAGGAACATTTATATCTATTCACCAAGATCTCCCTGGAATTTCTGGGCTTTCAGTTATGA 368

Qy 471 TGAGATGCTAGTTTGCACCTCTCTCGATGATAAATTTTATTTTGCAGAAACAGGCGCA 530

Db 369 TGAATGCTGATATGACCTTCTCATCTACATTTGATTTTATTTTAAAGAAAGACAGACA 428

Qy 531 GGAAAAGATCTATTATGTCTGGCTATTACAGGGCACCACCATGGCTTTTATTCATTTTC 590

Db 429 GAAGAAGCTACACTATGTGTGGCCATTCCCAAGGCACCACCATGTTGTTTGTGCTGCTTTTC 488

Qy 591 CACCATGCCAGAGCTGGCTCAGAAATCAAAATGTTTGTCTTTAGCACCACATAGCCAC 650

Db 489 TACCAATCCCACTGGCTGAAATTCGAAGTCTCCATGCAATAGCCCGCAGTCCGCAC 548

Qy 651 TGTTAAGCATGCAAAAGACCCCGGACCAAAATTTTGTCTGCGCAGATATGATGATCAA 710

Db 549 AGTGAAGCACACCCAGAGGCTGTTTAAACAACTTGCACTTATCTCTCTCTCTCTCAA 608

Qy 711 GCGATTTCTTTGGCAAAAGAAATTTCTGTATCAGACCAGATTTCTCAGACAACTTGTTAT 770

Db 609 GATTATATTTGGTAAACAAATGTTTACCCACACAAATTTTGTGAACAATTTCTTGTTGT 668

Qy 771 TTACCTTTTGGCCAGAGTGTGATTTGATCAGATTTGATTAATATCATGTTACTTCTGGG 830

Db 669 TGAAGTGTCTCTCTGTGAGACACTGGATGTCCTTTGTAAGAATGCTTTGTTGCCATTAC 728

Qy 831 TGGATTCACACCAACAAATATGAACATGAGCCGAGCAAGTGTATATGCTCTCCACACTCT 890

Db 729 TGGAGCTGACATTAATAAATTTCAACATGAGTCGTTAGATGTGTATGTAGCACATAATCC 768

Qy 891 TGCTGGAACTCTCTGCAAAATATTCTACACTGAGCCAGGCAAGTGAATTTCTGGTGAAT 950

Db 789 AGCAGGAGCTTCTGTTCAAAACATCTCCACTGAGACAGGCTATTAAAGTCTGGGAAAT 848

Qy 951 CCGGGCATTTGACTGGGGAGTGAGACCAAAATCTGAAAAATGCAATCAGCAACTCC 1010

Db 849 CCAAGCTTTTGTACTGGGAGGCTCAGTTGAGACCTAATGCATTAATCAGCCACACC 908

Qy 1011 TGTAAAGTACAGAGTCAGAGATATGACGGTCCCTPACAGCAATGTGGACAGAGTCA 1070

Db 909 TCCCATCTACAATTTAAACAGCCATGAATGTCCCAATTCAGATGTGAGTGTGCGCAAGA 968

Qy 1071 CTGGCTTTCAATCCAGAGAGCTGAAATGCTCTCTCTGAGTGACCACTCATCTA 1130

Db 969 CTGTGTGCTGACCTCAGGATGTTGACCTTTTCTTTTCTTTTCTTTTCTTTTCTTTTCT 1028

Qy 1131 CCATAAGAAATATCTCTGAAATGGGCTCAGTGGATTTTCATCTGGGGTTTGGATGCTCTCA 1190

Db 1029 CCACAGGAAATTTCCAAATTAATCATCTGGACTTTATCTGGCAATGATGACCTCA 1088

Qy 1191 CCGTATGTACAAATGAATCATCTCATCTGATGACGAGGAGAGA 1234

Db 1089 AGAAGTTTACAATGAATTTATTTCTTTGATGGCAAAAGACAAA 1132

RESULT 8

US-09-186-489-5

; Sequence 5, Application US/09186489

; Patent No. 6375947

; GENERAL INFORMATION:

; APPLICANT: Bolen, Paul L.

; APPLICANT: Cihak, Paul L.

; APPLICANT: Scharpf Jr., Lewis G.

; TITLE OF INVENTION: Purified Recombinant Kid Pregastric Esterase, and

; TITLE OF INVENTION: Processes for its Production and Use

; FILE REFERENCE: 5499/3

; CURRENT APPLICATION NUMBER: US/09/186,489

; CURRENT FILING DATE: 1998-11-05

; NUMBER OF SEQ ID NOS: 7

; SOFTWARE: PatentIn Ver. 2.0

; SEQ ID NO 5

; LENGTH: 8324

; TYPE: DNA

; ORGANISM: Yeast YE-1 expression vector

US-09-186-489-5

Query Match 35.4%; Score 449.8; DB 4; Length 8324;

Best Local Similarity 62.7%; Pred. No. 3.9e-135; Indels 0; Gaps 0;

Matches 700; Conservative 0; Mismatches 417; Indels 0; Gaps 0;

Qy 118 GCTGTGACCCAGAGCAATTCATGAATATTAGTGAATATCAATCAAGGTATCCCC 177

Db 1547 GCTAAGAACCCCTGAAGCCAGTATGAATGTGAGTCAGATGATTTCTTCTGGGGCTACCCA 1606

Qy 178 TGTGAGGAATATGAAGTGCAGCTGAAGATGGGTATATCTTCTCTTAAACAGATTCCT 237

Db 1607 AGTGAGATGCATAAGATTATACTGAGATGGGTATATCTTCTCAGGTCATTCGATTCCT 1666

Qy 238 CGAGGCTAGTGCACCTACCAAGACAGGTTCCAGGCTGTGGTGTACTGACAGATGGC 297

Db 1667 CATGGAAGAATGATGCTAATCATTTAGGTCAAGACCTGTTGTTGTTCTGCAGCATGGT 1756

298 CTAGTTGAGGTGCTAGCACTGATTTCCAACTGCGCCCAACATAGCTGGGCTTCATT 357
 1727 CTTCTTGCTCAGCTACAACTGATTTCCAACTTCCAAACACAGCTGGGCTTCCTC 1786
 358 CTGGCAGATGCTGTTTGAAGTGTGGATGGGAAACAGCAGGGGAAAGCGCTGCTCGA 417
 1787 CTGGCAGATGCTGTTTGAAGTGTGGATGGGAAACAGCAGGGGAAAGCGCTGCTCGA 1846
 418 AAACAACAGACACTCTCCATGACCAAGATGATTTCTGGGCTTTCAGTTATGATGATG 477
 1847 GAACATTATATCTATCCAGACATCCCTGAAATCTGGGCTTTCAGTTATGATGATG 1906
 478 GCTAGTTTGAAGTCTCTGCTGAGTATAAATTTTATTTTGCAGAAAGCGGCGCAAGAAAG 537
 1907 GCTGAATATGACCTTCCATCTACATTTGATTTTCAATTTTAAAGAGAACAGGACAGAG 1966
 538 ATCTATTATGCTGCTATTCAGAGGACCAACCATGAGCTTATTTGCAATTTTCCACCATG 597
 1967 CTACACTATGTTGGCCTATCCCAAGGACCAACCATGAGTTGTTTGTGGCTTTCACCAAT 2026
 598 CCAGAGCTGGCTCAGAAATCAAAATGATTTTGTCTTTAGCACCCATAGCCACTGTTAAG 657
 2027 CCCACACTGGCTGAAATTCGAAGTCTTCCATGATTTAGCCCCAGTCCGACAGTGAAG 2086
 658 CATGCAAAAGCCCGGAGACCAATTTTGTGCTGCGCAGATATGATGATCAAGGATG 717
 2087 CACACCCAGAGCTGTTTAAACAACTTGCACTTATTCCTCTCTTCAAGATATA 2146
 718 TTGCGCAAAAGAAATTTCTGATCAGACCAAGATTTCTCAGACAACTTGTATTACCTT 777
 2147 TTGCTGTAACAAATGTTCTACCCACAAATTTTGTGAACTTTCTTGTGTTGAAGT 2206
 778 TGTGCGCAGGTGATTTCTGATCAGATTTGTAGTAAATCATGTTACTTCTGGGTGATTC 837
 2207 TGCTCTGCTGAGACACTGATGCTCTTTGTAAGATGCTTTGTTGCCATTTACTGGAGCT 2266
 838 AACACCAACATATGAACATGAGCGAGCAGTATGATGCTGCCCCACACTTCTGCTGGA 897
 2267 GACATATAAACTTCAACATGATGCTGTTAGATGTTGATGATGATGATGATGATGATG 2326
 898 ACATCTGTGCAAAATTTCTACCTGGAGCCAGGACGATGAAATTTCTGTGAACTCCGGCA 957
 2327 GCTTCTGTTCAAAACATCTCCCTGAGAGCAGGCTATTAAGTCTGGGAAATTTCAAGCT 2386
 958 TTGAGTGGGAGTGAACCAAAATCTGGAATAATGCAATCAGCAGCACTCTCTGTAAG 1017
 2387 TTGAGTGGGAGCTCAGTTGAGAACCTAATGCAATTAATTAATCAGCCCACTCCCATC 2446
 1018 TACAGAGTCAGAGATATGACGCTCCCTACAGCAATGTCGACAGGAGGTGAGCTGCTT 1077
 2447 TACAATTTAAAGCCATGATGCTCCCAATTTGCAATGAGTATGAGTGTGGCCAGACCTGTTG 2506
 1078 TCAATTCAGAGAGCGTGAATGCTCTCTGAGGTGAGCAACCTCATCTACCAATAG 1137
 2507 GCTGACCTCTCAGGATGTTGACCTTTGCTTTCAAACTCTCTAATCTCATTCACCAAG 2566
 1138 ATATTTCTGATGGCTCAGGTGATTTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 1197
 2567 GAATTTCCAAATTAACATCATCTGAGCTTTATCTGGGCAATGAGTATGACCTTCAAGAGTT 2626
 1198 TACAATGAATCATCATCTGATGACGAGGAGGAGA 1234
 2627 TACAATGAATTTTCTTTGATGCAAAAGACAAA 2663

RESULT 9

US-10-043-665B-5
 ; Sequence 5, Application US/10043665B
 ; Patent No. 6582948
 ; GENERAL INFORMATION:
 ; APPLICANT: Bolen, Paul L.
 ; APPLICANT: Cihak, Paul, L.

APPLICANT: Scharpf Jr., Lewis G.
 TITLE OF INVENTION: Recombinant Kid Progastric Esterase and Methods for Its
 TITLE OF INVENTION: Production and Use
 FILE REFERENCE: IFF-0009
 CURRENT APPLICATION NUMBER: US/10/043,665B
 CURRENT FILING DATE: 2002-09-24
 PRIOR APPLICATION NUMBER: US 09/186,489
 PRIOR FILING DATE: 1998-11-05
 NUMBER OF SEQ ID NOS: 7
 SOFTWARE: PatentIn Ver. 2.0
 SEQ ID NO 5
 LENGTH: 8324
 TYPE: DNA
 ORGANISM: Yeast YE-1 expression vector
 US-10-043-665B-5

Query Match 35.4%; Score 449.8; DB 4; Length 8324;
 Best Local Similarity 62.7%; Pred. No. 3.9e-135;
 Matches 700; Conservative 0; Mismatches 417; Indels 0; Gaps 0;

118 GCTGTGAGCCCAAGACATTCATGAATATTAGTGAATCATCCAACTCAAGGCTATCCG 177
 1547 GCTAAGAACCTTGAGCCAGTATGAATGTGAGTCAGATGATTTCTTCTGGGCTACCCA 1606
 178 TGTGAGGAATATGAAGTCCAACTGAAGATGGGTATATCTTTCTGTTAAAGGATTCCT 237
 1607 AGTGAGATGATAAAGTTATAAATGAGATGCTATATCTTTTCAAGTCTATCGGATTCCT 1666
 238 CGAGGCTAGTGCACCTAAGAGACAGGTTCCAGGCTGTGTTTCTACTGAGGATGCG 297
 1667 CATGGAAGATGATGCTAATCATTTAGTCTAGAGACCTGTTGTTTCTGAGGATGCT 1726
 298 CTAGTTGAGGTGCTAGCACTGATTTCCAACTGCCCCAACAAATAGCTGGGCTTCATT 357
 1727 CTTCTTGCTCAGCTACAACTGGAATTTCCAACTTCCCAACCAAGCTGGGCTTCCTC 1786
 358 CTGGCAGATGCTGTTTGAAGTGTGGATGGGAAACAGCAGGGGAAAGCGCTGCTCGA 417
 1787 CTGGCAGATGCTGTTTGAAGTGTGGATGGGAAACAGCAGGGGAAAGCGCTGCTCGA 1846
 418 AAACAACAGACACTCTCCATGACCAAGATGATTTCTGGGCTTTCAGTTATGATGATG 477
 1847 GAACATTATATCTATCCAGACATCCCTGAAATTTCTGGGCTTTCAGCTTTCAGAAATG 1906
 478 GCTAGGTTTGAAGTCTCTGCTGAGTATAAATTTTATTTTGCAGAAACAGGCGCAAGAAAG 537
 1907 GCTGAATATGAGCTTCCATCTAATTTGATTTTCACTTTAAAGAGAACAGGACAGAG 1966
 538 ATCTATTATGCTGCTGCTATTCAGAGGACCAACCATGAGGCTTTCAGCTTTTCCACCATG 597
 1967 CTACACTATGTTGGCCTATTCAGAGGACCAACATGTTGTTTGTGGCTTTCACCAAT 2026
 598 CCAGAGCTGGCTCAGAAATCAAAATGATTTTGTCTTTAGCAACCCATAGCCACTGTTAAG 657
 2027 CCACACTGCTGAGAAATTCGAAGTCTTCCATGATTTAGCCCCAGTCCGACAGTGAAG 2086
 658 CATGCAAAAGCCCGGAGACCAATTTTGTGCTGCGCAGATATGATGATCAAGGATG 717
 2087 CACACCCAGAGCTGTTTAAACAACTTGCACTTATTCCTCTCTTCAAGATATA 2146
 718 TTGCGCAAAAGAAATTTCTGATCAGACCAAGATTTCTCAGACAACTTGTATTACCTT 777
 2147 TTGCTGTAACAAATGTTCTACCCACAAATTTTGTGAACTTTCTTGTGTTGAAGT 2206
 778 TGTGCGCAGGTGATTTCTGATCAGATTTGTAGTAAATCATGTTACTTCTGGGTGATTC 837
 2207 TGCTCTGCTGAGACACTGATGCTCTTTGTAAGATGCTTTGTTGCCATTTACTGGAGCT 2266
 838 AACACCAACATATGAACATGAGCGAGCAGTATGATGCTGCCCCACACTTCTGCTGGA 897
 2267 GACATATAAACTTCAACATGATGCTGTTAGATGTTGATGATGATGATGATGATGATG 2326
 898 ACATCTGTGCAAAATTTCTACCTGGAGCCAGGACGATGAAATTTCTGTGAACTCCGGCA 957

Db 2327 GCCTCTGCTCAAAACATCTCTCCACTGGAGACAGCTATTAAAGTCCTGGAAATCCAGCT 2386
 Qy 958 TTTGACTGGGGAGTGAGACAAATAATCTGGAAAAATGCATCAGCCAACTCTCTGTAAG 1017
 Db 2387 TTTGACTGGGGAGCTCAGTTGAGAACCTTAATGCATTTAATCAGCCCACTCTCCATC 2446
 Qy 1018 TACAGACTCAGAGATATGACGGTCCCTACAGCAATGTGGACAGAGGTGAGGACTGGCTT 1077
 Db 2447 TACAAATTAACAGCCATGATGTCCTCAATTTGAGATGAGTGTCTGGCCAGACCTGTTG 2506
 Qy 1078 TCAATCCAGAGACGTGAAAAATGCTGCTCTCTGAGGTGACCACTCATCTACCAATAG 1137
 Db 2507 GCTGACCTCAGGATGTTGACCTTTTGTCTTCAAAACTCTCTAATCTCAATCACCACAAG 2566
 Qy 1138 AATATTCCTGAATGGGCTCAGGTGATTTTCATCTGGGTTTGTGATGCTCTCACGATG 1197
 Db 2567 GAATTCCTCAATTAACAATCATCTGGACTTTTATCTGGCAATGATGCACTCAAGATT 2626
 Qy 1198 TACAATGAATTCATCCATCTGATGACAGCAGGAGAGA 1234
 Db 2627 TACAATGAATTTATTTCTTTGATGGCAAAAGACAAA 2663

RESULT 10

US-08-227-108-2
 ; Sequence 2, Application US/08227108
 ; Patent No. 5807726
 ; GENERAL INFORMATION:
 ; APPLICANT: Blanchard, Claire
 ; APPLICANT: Benicourt, Claude
 ; APPLICANT: Julien, Jean-Louis
 ; TITLE OF INVENTION: Recombinant Dog Gastric Lipase
 ; NUMBER OF SEQUENCES: 21
 ; CORRESPONDENCE ADDRESS:
 ; ADDRESSEE: Pennie & Edmonds
 ; STREET: 1155 Avenue of the Americas
 ; CITY: New York
 ; STATE: New York
 ; COUNTRY: U.S.A.
 ; ZIP: 10036
 ; COMPUTER READABLE FORM:
 ; MEDIUM TYPE: Floppy disk
 ; COMPUTER: IBM PC compatible
 ; OPERATING SYSTEM: PC-DOS/MS-DOS
 ; SOFTWARE: Patent In Release #1.0, Version #1.25
 ; CURRENT APPLICATION DATA:
 ; APPLICATION NUMBER: US/08/227,108
 ; FILING DATE: 03-APR-1994
 ; CLASSIFICATION: 435
 ; ATTORNEY/AGENT INFORMATION:
 ; NAME: Fanucci, Allan A.
 ; REGISTRATION NUMBER: 30,256
 ; REFERENCE/DOCKET NUMBER: 7620-033
 ; TELECOMMUNICATION INFORMATION:
 ; TELEPHONE: 212 790-9090
 ; TELEFAX: 212 869-8864/9741
 ; TELEX: 66141 PENNIE
 ; INFORMATION FOR SEQ ID NO: 2:
 ; SEQUENCE CHARACTERISTICS:
 ; LENGTH: 1137 base pairs
 ; TYPE: nucleic acid
 ; STRANDEDNESS: double
 ; TOPOLOGY: linear
 ; MOLECULE TYPE: cDNA to mRNA
 ; FEATURE:
 ; NAME/KEY: CDS
 ; LOCATION: 1..1137
 ; US-08-227-108-2

Query Match 35.0%; Score 444.2; DB 1; Length 1137;
 Best Local Similarity 62.8%; Pred. No. 8e-134;
 Matches 689; Conservative 0; Mismatches 408; Indels 0; Gaps 0;

Qy 125 ACCCAGAGCATTCATGAATATTAGTGAATCATCAACATCAAGGCTATCCCTGTGAGG 184
 Db 26 ACCCTGAAAGTACCATGATATATAGTCAATGATCACTACTGGGATACCAAGCTGAGG 85
 Qy 185 AATATGAAGTCGCAACTGAAGATGGGTATATCTCTTTCTGTAAACAGAGATTCTCTGAGGCC 244
 Db 86 AATATGAAGTTGTGACCGAAGACGGTTATATCTCTGGGATCGACAGAAATTCCTTATGGGA 145
 Qy 245 TAGTGCACCTTAAGAGACAGAGGTTCCAGGCTGTGGTGTACTGCAGCATGGCTAGCTTG 304
 Db 146 GGAATAATTCAGAGAAATATAGGCGGAGACCTGTTCATTTTTCACACAGGTTTGTCTCG 205
 Qy 305 GAGGTGCTAGCAACTGGATTTCCAACCTGCCCCCAACAATAGCCTGGGCTTCACTTCTGSCAG 364
 Db 206 CATCAGGCCACAACTGGATCTCCAACTGCCCCCAACAACAGCCTGGCCTTCATCTGCGCG 265
 Qy 365 ATGCTGTTTTGAGGTGTGATGGGAAACAGCAGGCGGAAAGCGCTGCTCGAAAAACACA 424
 Db 266 AGCGGGGTACGAGCTGTGGCTGGGAAACAGCAGGCGGAAACACCTGGGCGCAGGAGATC 325
 Qy 425 AGACACTCTCCATAGACCAAGATGAGTTCTGGGCTTTTCAGTTATGATGAGATGCTAGGT 484
 Db 326 TGTACTACTCGCCCGACTCCGTCGAATCTGGGCTTTTCAGCTTTGACGAGATGCTAAAT 385
 Qy 485 TTGACCTTCTGAGTGATGATAAATTTTATTTTGCAGAAAAACGGGCCAGGAAAAAGATCTATT 544
 Db 386 ATGACCTTTCCGCGCCACCATTTGACTTTCATTTTGAAGAAAAACGGGACAGCAAGCTACCT 445
 Qy 545 ATGCTGGCTATTTCAGGCGCACCACTGGGCTTTTATTTGCAATTTTCCACCATGCGCAGAGC 604
 Db 446 AGTTGGCCATTCAGGCGCACCACTGGGCTTTTATTCGCTTTTCCACCAATCCCAAGC 505
 Qy 605 TGGCTCAGAAAAACAAATGATTTTGTGTAGCAACCACTAGCCTGTTTAAAGCATGCA 664
 Db 506 TGGCGAAACGGATCAAAACCTTCTATGCAATAGCTCCGTTCCGACCGCTGAAGTACACCG 565
 Qy 665 AAAGCCCGGAGCAAAATTTTGTGTGCGCAGATATGATGATCAAGGATTTGTTGGCA 724
 Db 566 AAACCTTGTAAACAAATCATGCTGCTCTTCTGCTCTTCAAGCTTATATTTGGAA 625
 Qy 725 AAAAAAGATTTCTGTATCAGACCAAGATTTCTCAGACAACTTTGTTTATTTTACCTTTGCGCC 784
 Db 626 ACAAAATATTTACCCACACCACTCTTTGATCAATTTCTGCCACCGAGGATAGTCTCCC 685
 Qy 785 AGTGATTTCTGATCAGATTTGTAGTATATATCATGTTACTTCTGGTGGATTCAACACCA 844
 Db 686 GCGAGACGGTGGATCTCTCTGACGAAACGCGCTGTTTATCATTTGTGAGATTGACACTA 745
 Qy 845 ACAATATGAACATGAGCGCAGCAAGTGTATATGCTGCCCACTCTTGTCTGGAACATCTG 904
 Db 746 TGAATTTGACATGATGCTGCTGGATGTATCTGTCAATATCCAGCAGGAAACATCGG 805
 Qy 905 TCGAAAAATTTCTACACTGGAGCAGGAGTGAATTTGTTGTAATCTCGGCGCATTTGACT 964
 Db 806 TTCAGAACGCTGCTCCACTGCTCCAGGCTGTTAAGTCTGGGAAGTTCCAAAGCTTTTGACT 865
 Qy 965 GGGGAGTGAGACCAAAATCTGGAATAATGCAATCAGCAACTCTCTTAAGGTACAGAG 1024
 Db 866 GGGGAGGCCAGTTTCAAGAACATGATGCACTATCATCAGAGCATGCTCTCTACTACAC 925
 Qy 1025 TCAGAGATATGACGGTCCCTACAGCAATGTGGACAGGAGGTGAGGCTGGTTTCAATC 1084
 Db 926 TGACAGACATGATGCTGCAATCGCAGTGTGGAACGGTGGCAACGCACTGCTGGCGGACC 985
 Qy 1085 CAGAGACGTGAAATGCTGCTCTGAGGTGACCACTCATCTACCATAGGATATTC 1144
 Db 986 CTCACGATGTTGACCTTTTGTCTTCAAGCTCCCCAATCTCATTTTACCACAGGAGATTC 1045
 Qy 1145 CTGAATGGGCTCAGCTGGATTTTCACTGGGTTTGGATGCTCCTCACCGTATGTACATG 1204
 Db 1046 CTCCTTACATCATCTGGACTTTTATCTGGGCCATGATGCTCCCTCAAGCGGTTTACATG 1105

QY 1205 AAATCATCCATCTGATG 1221
 DB 1106 AAATGTTTCCATGATG 1122

RESULT 11

US-09-073-674-2
 ; Sequence 2, Application US/09073674
 ; Patent No. 5958189
 ; GENERAL INFORMATION:
 ; APPLICANT: Blanchard, Claire
 ; APPLICANT: Benicourt, Claude
 ; APPLICANT: Junien, Jean-Louis
 ; TITLE OF INVENTION: Recombinant Dog Gastric Lipase
 ; NUMBER OF SEQUENCES: 21
 ; CORRESPONDENCE ADDRESS:
 ; ADDRESSEE: Warner-Lambert Company
 ; STREET: 2800 Plymouth Road
 ; CITY: Ann Arbor
 ; STATE: Michigan
 ; COUNTRY: U.S.A.
 ; ZIP: 48105

COMPUTER READABLE FORM:
 ; MEDIUM TYPE: Floppy disk
 ; COMPUTER: IBM PC compatible
 ; OPERATING SYSTEM: PC-DOS/MS-DOS
 ; SOFTWARE: Patent In Release #1.0, Version #1.25
 ; CURRENT APPLICATION DATA:
 ; APPLICATION NUMBER: US/09/073,674
 ; FILING DATE:
 ; CLASSIFICATION:
 ; ATTORNEY/AGENT INFORMATION:
 ; NAME: Crissey, Todd M.
 ; REGISTRATION NUMBER: 37,807
 ; REFERENCE/DOCKET NUMBER: 5072-D1-66-TWC
 ; TELECOMMUNICATION INFORMATION:
 ; TELEPHONE: 734 622-7530
 ; TELEFAX: 734 622-1553
 ; TELEX:
 ; INFORMATION FOR SEQ ID NO: 2:
 ; SEQUENCE CHARACTERISTICS:
 ; LENGTH: 1137 base pairs
 ; TYPE: nucleic acid
 ; STRANDEDNESS: double
 ; TOPOLOGY: linear
 ; MOLECULE TYPE: cDNA to mRNA
 ; FEATURE:
 ; NAME/KEY: CDS
 ; LOCATION: 1..1137

US-09-073-674-2

Query Match 35.0%; Score 444.2; DB 2; Length 1137;
 Best Local Similarity 62.8%; Pred. No. Be-134;
 Matches 689; Conservative 0; Mismatches 408; Indels 0; Gaps 0;

QY 125 ACCGAGGACATTCATGATTAAGTGAATCCTCAACATCCAGGCTATCCCTGTGAGG 184
 DB 26 ACCCTGGAAGTGAACATGAATTAAGTCAGATGATCCTACTGGGGATACCCAGCTGAGG 85
 QY 185 AATATGAAGTCGCAACTGAAGATGGGTATATCCCTTCTGTTAAACAGGATTCCTCGAGGCC 244
 DB 86 AATATGAAGTTGTGACCGAAGACGGTATATATCCCTTGGGATCGACAGAATTCCTTATGGGA 145
 QY 245 TAGTGCACCTTAAGAGACAGGTTCCAGGCCCTGGTGTACTGAGCATGGCCCTAGTTG 304
 DB 146 GGAATAATTCAGAGATATAGCCGGAGACCTGTTGCAATTTTTCACACAGGTTTGTCTG 205
 QY 305 GAGGTGCTAGCAACTGGATTTCCAACTGCCCCCAACATAGCTGGGCTTCACTCTGGCAG 364
 DB 206 CATCAGCCACAACCTGGATCTCCACCTGCCCCCAACAGCTGGCCTTCACTCTGGCAG 265
 QY 365 ATGCTGGTTTGAAGTGGATGGGGAACAGCAGGGGAAACCCCTGGTCTCGAATAACACA 424

DB 266 ACGCCGGGTACGACGTGTGGCTGGGGAAACAGCAGGGGCAACACCTGGGCCAGGAGATC 325
 QY 425 AGACACTCTCCATAGACCAAGATGAGTTCTGGGCTTTTCAAGTTATGATGATGAGTGGCTAGGT 484
 DB 326 TGTACTACTCGCCGACCTCGGTGCAATTTCTGGGCTTTTCAAGCTTTTACGAGATGGCTAAAT 385
 QY 485 TTGACCTTCTCTGAGTGATTAATTTTTCAGAAAACGGGCCAGGAAAAGATCTATT 544
 DB 386 ATGACCTTCCCGCCACCATTTGATCTTCAATTTGAAGAAAACGGGACACGACAGCTACACT 445
 QY 545 ATGTCGGCTATTTCAGAGGACACCATGGGCTTTTATGCAATTTTCCACCATGCGCAGAC 604
 DB 446 ACCTTGGCCATTCGCCAGGACACCATTTGTTTATGCGCTTTTCCACCAATCCCAAGC 505
 QY 605 TGGCTCAGAAAATCAAAATGTTATTTTGTAGTCCACCATAGCACTGTTAGCATGCA 564
 DB 506 TGGCGAAACGGATCAAAACCTTCTATGCAATTAGTCCCGTTCACCGGTGAAGTACACCG 565
 QY 665 AAAGCCCCGGGACCAAAATTTTGTCTGCCAGATATGATCAAGGATTTGTTTGGCA 724
 DB 566 AAACCTGTAAACAAACTCATGCTCGTCCCTTCTGTTCTCTTCAAGCTTATTTGGAA 625
 QY 725 AAAAAGAAATTTCTGATCAGACAGATTTCTCAGACAACTGTTATTTACCTTTTGGGC 784
 DB 626 AAAAAATATTTTACCCACACCACTTCTTTTATGATCAATTTTCCGCCACCGAGGTATGCTCC 685
 QY 785 AGTGATTTCTGATCAGATTTGTTAGTAAATATCATGTTACTTCTGGGTGATTCACACCA 844
 DB 686 GCGAGACGGTGGATCTCCTCTGAGCAACGCCCTGTTTATCAATTTGTTGATTTGACACTA 745
 QY 845 ACAATATGAACATGAGCCGAGCAAGTGTATATGTCGCCCACTCTTGTCTGAAACATCTG 904
 DB 746 TGAACCTTGAACATGAGTCTGTTGATGTATCTGTACATATCCAGCAGGAAACATCGG 805
 QY 905 TGCATAATATTCTACACTGAGCCGACGAGTGAATTTCTGTTGAACTCCGGGCAATTTGACT 964
 DB 806 TTCAAGACGTGCTCCACTGCTCCAGGCTGTTAAGTCTGGGAAGTTTCAAGCTTTTACT 865
 QY 965 GGGGAGTGAAGACCAAAATCTGAAAATCAATCAAGCAACTCTCTGTAAGGTACAGAG 1024
 DB 866 GGGGAAGCCAGTTTCAAGACATGATGATCATCATCAGAGCATGCTCTCTACTACAAC 925
 QY 1025 TCAGATATGAGGTCCTTACAGCAATGTCGACAGGAGTCAAGCATGGCTTTCAATC 1084
 DB 926 TGACAGACATGATGTGCCAATCGCAGTGTGAAACGGTGGCAACGACTTGTCTGGCCGACC 985
 QY 1085 CAGAAGACGTGAAAATGCTGCTCTCTGAGGTGACCAACCTCATCTACCATAGAATATTC 1144
 DB 986 CTCACGATGTTGACCTTTTGTCTTCAAGCTCCCAATCTCATTTACACAGGAAGATTC 1045
 QY 1145 CTGAATGGGCTCAGGTTGATTTTCACTGGGGTTTGGATGCTCTCTCACCGTATGTACAATG 1204
 DB 1046 CTCCTTACAATCACTTGGACTTTATCTGGGCGCATGATGATGCCCTCAAGGGGTTTACAATG 1105
 QY 1205 AATCATCCATCTGATG 1221
 DB 1106 AAATGTTTCCATGATG 1122

RESULT 12

US-08-227-108-4
 ; Sequence 4, Application US/08227108
 ; Patent No. 5807726
 ; GENERAL INFORMATION:
 ; APPLICANT: Blanchard, Claire
 ; APPLICANT: Benicourt, Claude
 ; APPLICANT: Junien, Jean-Louis
 ; TITLE OF INVENTION: Recombinant Dog Gastric Lipase
 ; NUMBER OF SEQUENCES: 21
 ; CORRESPONDENCE ADDRESS:
 ; ADDRESSEE: Pennie & Edmonds
 ; STREET: 1155 Avenue of the Americas
 ; CITY: New York

STATE: New York
COUNTRY: U.S.A.
ZIP: 10036
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patent in Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/227.108
FILING DATE: 03-APR-1994
CLASSIFICATION: 435
ATTORNEY/AGENT INFORMATION:
NAME: Famucci, Allan A.
REGISTRATION NUMBER: 30,256
REFERENCE/DOCKET NUMBER: 7620-033
TELECOMMUNICATION INFORMATION:
TELEPHONE: 212 790-9090
TELEFAX: 212 869-8864/9741
INFORMATION FOR SEQ ID NO: 4:
SEQUENCE CHARACTERISTICS:
LENGTH: 1140 base pairs
TYPE: nucleic acid
STRANDEDNESS: double
TOPOLOGY: linear
MOLECULE TYPE: cDNA to mRNA
FEATURE:
NAME/KEY: CDS
LOCATION: 1..1140
US-08-227-108-4

Query Match 35.0%; Score 444.2; DB 1; Length 1140;
Best Local Similarity 62.8%; Pred. No. Se-134;
Matches 689; Conservative 0; Mismatches 408; Indels 0; Gaps 0;

QY	125	ACCCAGAGCATTCATGAATATTAGTGAATATCATCAACATCAAGGCTATCCCTGTGAGG	184
DB	29	ACCTGAGTGNCCATGATATAGTCAATGATCACTACTGGGATACCCAGCTGAGG	88
QY	185	AATATGAAGTCGCAACTGAAGATGGGTATATCCCTTTCGTTAAACAGGATTCCTCGAGCC	244
DB	89	AATATGAAGTTGTGACCGAAGCGTTATATCCCTGGGATCGACAGAAATTCCTTATGGGA	148
QY	245	TAGTCAACCTTAAGAGACAGGTTCCAGGCTGTGGTACTACGATGCGCTAGTGTG	304
DB	149	GGAAAAATTCAGAGAAATAGCGGAGACCTGTTCATTTTGGCAACGCTTGTCTG	208
QY	305	GAGGTGCTAGCAACTGGATTTCCAACTGCCCCAACAATAGCTGGGCTTTCATTTCTGGCAG	364
DB	209	CATCAGCCACAACTGGATCTCCAACTGCCCAACAACAGCTGSCCTTCATCTCTGGCCG	268
QY	365	ATGCTGGTTTTCAGCTGTGGATGGGAAACAGAGGGGAAACGCTGTCTCGAAACACA	424
DB	269	ACGCGGGGTACACGTGTGGTGGGGAACAGAGGGGCAACACCTGGGCGCAGGAGGAATC	328
QY	425	AGACACTCTCCATAGACCAAGATGATGTTCTGGGCTTTCAGTTATGATGAGATGCTAGGT	484
DB	329	TGTACTACTGCGCCGACTCCGTCGAATTTCTGGCTTTCAGCTTTGACGAGATGCTAAAT	388
QY	485	TTGACCTTCTGCACTGATAAATTTATTTTTCAGAAAAACGGGCGCAGGAAAGATCTATT	544
DB	389	ATGACCTTCCCGCCACCATTCATCTTGAAGAAACGGGACAGGACAGCAAGCTACACT	448
QY	545	ATGTCGGCTATTTCAGAGGCAACCATGGGCTTATTCATTTTCCACCATGCGAGGC	604
DB	449	ACGTTGGCCATTCACAGGCGACCAATGTTTTCATGCTCTTTCACCAATCCCAAGC	508
QY	605	TGGCTCAGAAAAATCAAAATGTTATTTTGTAGTACCCATAGCCACTGTTAAGCATGCAA	664
DB	509	TGGCGAAACGGATCAAAACCTTCTATGATTAGCTCCCGTTGCGCAACGCTGAAGTACACCG	568
QY	665	ANAGCCCCGGGACCAAAATTTTGTGTCGCCAGATATGATCAAGGATGTTTGCA	724

DB	569	AAACCCCTGTTAAACAAACTCATGCTCGTCCCTCGTCTCTTCAAGCTTATATTGGAA	628
QY	725	AAAAAGAAATTTCTGTATCAGACCAGATTCTCAGACAACTTGTATTACCTTTGTGGCC	784
DB	629	ACAAAATATTTCACCCACACACTTCTTTGATCAATTTCTGCCACCGAGTATGTCTCCC	688
QY	785	AGGTGATTCTTGATCAGATTTGTAGTAAATCATGTTACTTCTCTGGGTGATTCACACCA	844
DB	689	GGGAGACGCTGATCTCTCTGCAACAAGCCCTGTTTATCATTTGTGGATTGACACTA	748
QY	845	ACAAATATGAACATGAGCCGAGCAAGTATATGTCGCCACACTCTTCTGCTGAACATCTG	904
DB	749	TGAACCTTGAACATGATGCTGCTGGATGTGTATCTGTACATAATCCAGCAGGAACATCG	808
QY	905	TGCAAAATATTCTACACTGGAGCCAGGAGTGAATTTCTGTGCAACTCCCGGCACTTTGACT	964
DB	809	TTCAAGACGCTGCTCCACTGCTCCAGGCTGTTAGTCTGGGAAGTTCCAAAGCTTTGACT	868
QY	965	GGGGAGTGAGACCAAAATCTGGAATAATGCAATCAGCAACTCTCTGTAAGGTACAGAG	1024
DB	869	GGGGAAGCCAGTTTCAGAAACATGATGCACTATCATCAGAGCATGCTCTCTACTACACC	928
QY	1025	TCAGAGATATGACGCTCCCTACAGCAATGTGACAGAGGTCAGGACTGGCTTTCAAAATC	1084
DB	929	TGACAGACATGATGTGCAATCGCAGTGTGGAAGGTGGCAACGACTTGTGTCGCGGACC	988
QY	1085	CAGAAAGCGTGAAATGCTGCTCTGAGGTGACCAACTCATCTTACATGAAGATATTC	1144
DB	989	CTCAGCATGTTGACCTTTTGTCTTCCAAAGCTCCCAATCTCATTTACCAAGGAATTC	1048
QY	1145	CTGAATGGCTCAGCTGGATTTTCATCTGGGTTTGGATGCTCTCCACCGTATGTACAATG	1204
DB	1049	CTCCTTACAAATCAGTTGACTTTTATCTGGGCAATGATGCTCTCAAGCGGTTTACAATG	1108
QY	1205	AAATCATCCATCTGATG 1221	
DB	1109	AAATGTTTCCATGATG 1125	

RESULT 13
US-09-073-674-4
Sequence 4, Application US/09073674
Patent No. 5988189
GENERAL INFORMATION:
APPLICANT: Blanchard, Claire
APPLICANT: Benicourt, Claude
APPLICANT: Junier, Jean-Louis
TITLE OF INVENTION: Recombinant Dog Gastric Lipase
NUMBER OF SEQUENCES: 21
CORRESPONDENCE ADDRESS:
ADDRESSEE: Warner-Lambert Company
STREET: 2800 Plymouth Road
CITY: Ann Arbor
STATE: Michigan
COUNTRY: U.S.A.
ZIP: 48105
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patent in Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/09/073,674
FILING DATE:
CLASSIFICATION:
ATTORNEY/AGENT INFORMATION:
NAME: Crissey, Todd M.
REGISTRATION NUMBER: 37,807
REFERENCE/DOCKET NUMBER: 5072-D1-66-TMC
TELECOMMUNICATION INFORMATION:
TELEPHONE: 734 622-7530
TELEFAX: 734 622-1553

86 AATATGAAGTTGTGACCGAGAGCGTTATATCTTGGGATCGACAGAAATCTTATGGGA 145
 245 TAGTGCAACCTAAGAAGACAGGTTCCAGGCTGTGTGTACTGACGATGGCTAGTTG 304
 146 GGAATAATCAGAGATATAGGCGGAGAGCTGTGCATTTTGGCAACAGGTTGCTCG 205
 305 GAGTGTAGCACTGATTTCAACCTGCCCAACATAGCTGGCTTCAATCTGGCAG 364
 206 CATCAGCCACAACTGATCTCCAACTGCCCAACACAGCTGGCTTCACTCTGGCG 265
 365 ATGCTGTTTGTAGCTGTGATGGGGAACAGCGGGGAAACGCTGTCTCGAAACACA 424
 266 ACGCGGGTAGAGCTGTGGGGAACAGAGGGGCAACACCTGGCGCAGGAGATC 325
 425 AGACACTCTCCATAGACCAAGATGATTTCTGGGCTTTCAGTTATGATGATGGCTAGT 484
 326 TGTAATCTGCGGCACTCCGTGCAATTTCTGGGCTTTCAGTTTGAAGATGGCTAAT 385
 485 TTGACCTTCTGCAAGTAACTTTATTTTGGCAAAACGGGCGCAGGAAAGATCTATT 544
 386 ATGACCTTCCGCGCACCACTGATTTCACTTTGAAGAAACGGGACAGGCAAGCTACAT 445
 545 ATGTCGGCTATTCCAGGGCACCACTGAGCTTTATGCAATTTCCACCTGCCAGAGC 604
 446 ACGTTGGCCCATCCAGGGCACCACTGATTTGTCATGCGCTTTTCCCAATCCCAAGC 505
 605 TGGCTCAGAAAAATCAAAATGATTTTCTTTAGCACCCATAGCACTGTTAAGCATGCAA 664
 506 TGGCGAAACGATCAAAACCTTCTATGATAGCTCCCGTTGCCACCGTGAAGTACACCG 565
 665 AAAGCCCGGGACCAATTTTGTGTCGCGCAATGATGATCAAGGATGTTTGGGA 724
 566 AAACCTTGTAAACAACTCATGCTGCTGCTTCTGTTCTTCAAGCTTATATTTGGAA 625
 725 AAAAAGAAATTTCTATCAGACAGATTTCTCAGACAACTTGTATTTTACCTTTGGGC 784
 526 ACAAATTTCTACCCACACCACTTCTTTGATCAATTTCTGCGCACCGAGGTATGCTCC 585
 785 AGGTGATTTCTGATCAGATTTGATGATTAATATCATGTTATCTTGGGTGATTCACCA 844
 686 GCGAGACGGTGGATCTCTCTGACGCAACCGCTTGTATCATTTGTGATTTGACACTA 745
 845 ACATATGAACATGAGCGGAGCAAGTGTATGTCGCCACACTCTTGTCTGGAACATCTG 904
 746 TGAATTTGAACATGAGTGTGCTGATGATGATCTGTCACATATCAGCAGGAAATCG 805
 905 TGCAAAATATTTCTACTGAGCGCAGGAGTGAATTTCTGTTGAATCTCGGCGCATTTGACT 964
 806 TCCAGACGTGCTCCACTGTTCCAGGCTGTTAAGTCTGGGAAGTTTCAAGCTTTTGA 865
 965 GGGGAGTGAGACCAAAATCTGAAAAATGCAATCAGCCAACTCTGTTAAGTACAGAG 1024
 866 GGGGAAGCCCAAGTTTCAGAAATGATGATCATCATCAGAGCATGCTCTCCCTACTCA 925
 1025 TCAGAGATATGACGGTCCCTACAGCAATGTGGACAGGAGTCAAGGCTGCTTCAAAATC 1084
 926 TGAAGACATGATGTGCAATCGAGTGTGGACAGGTTGCAAGCTTCTGCGCGGAC 985
 1085 CAGAAGAGCGTGAATGCTGCTCTGAGGTGACCAACCTCATCTACCATTAAGATATTC 1144
 986 CTCAGATGTTGACCTTTGCTTTCCAAAGCTCCCAATCTCATTTACCAAGGAGATTC 1045
 1145 CTGAATGGGCTCAGTGAATTTCTATGCGGGTTTGGATGCTCTCAAGCTATGATGATG 1204
 1046 CTCCTTAAATCACTTGAATTTTATCTGGGCGATGGAATGCGCTTCAAGCGGTTTACA 1105
 1205 AAATCATCCATCTGATG 1221
 1106 AAATGTTTCCATGATG 1122

RESULT 15

US-09-073-674-6
 ; Sequence 6, Application US/09073674
 ; Patent No. 598189
 ; GENERAL INFORMATION:
 ; APPLICANT: Blanchard, Claire
 ; APPLICANT: Benicourt, Claude
 ; APPLICANT: Junien, Jean-Louis
 ; TITLE OF INVENTION: Recombinant Dog Gastric Lipase
 ; NUMBER OF SEQUENCES: 21
 ; CORRESPONDENCE ADDRESS:
 ; ADDRESSEE: Warner-Lambert Company
 ; STREET: 2800 Plymouth Road
 ; CITY: Ann Arbor
 ; STATE: Michigan
 ; COUNTRY: U.S.A.
 ; ZIP: 48105
 ; COMPUTER READABLE FORM:
 ; MEDIUM TYPE: Floppy disk
 ; COMPUTER: IBM PC compatible
 ; OPERATING SYSTEM: PC-DOS/MS-DOS
 ; SOFTWARE: PatentIn Release #1.0, Version #1.25
 ; CURRENT APPLICATION DATA:
 ; APPLICATION NUMBER: US/09/073,674
 ; FILING DATE:
 ; CLASSIFICATION:
 ; ATTORNEY/AGENT INFORMATION:
 ; NAME: Cribsey, Todd M.
 ; REGISTRATION NUMBER: 37,807
 ; REFERENCE/DOCKET NUMBER: 5072-D1-66-TMC
 ; TELECOMMUNICATION INFORMATION:
 ; TELEPHONE: 734 622-7530
 ; TELEFAX: 734 622-1553
 ; TELEX:
 ; INFORMATION FOR SEQ ID NO: 6:
 ; SEQUENCE CHARACTERISTICS:
 ; LENGTH: 1146 base pairs
 ; TYPE: nucleic acid
 ; STRANDEDNESS: double
 ; TOPOLOGY: linear
 ; MOLECULE TYPE: cDNA to mRNA
 US-09-073-674-6

Query Match 35.0%; Score 444.2; DB 2; Length 1146;
 Best Local Similarity 62.8%; Pred. No. 8e-134;
 Matches 589; Conservative 0; Mismatches 408; Indels 0; Gaps 0;
 QY 125 ACCGAGAGCAATTCATGAATATTAGTGAATATCCATTCCTGTTAAACAGGATTCCTCGAGGCC 244
 DB 26 ACCCTGAAGTGACCAATGAATATAAGTCAGATGATCACCTACTGGGGATACCCAGCTGAGG 85
 QY 185 AATATGAAGTCGCAACTGAAGATGGGTATATCTCTGTTAAACAGGATTCCTCGAGGCC 244
 DB 86 AATATGAAGTTGTGACCGAGACGTTATATCTCTGGATCGACAGATTCCTTATGGGA 145
 QY 245 TAGTGCAACCTAAGAAGACAGGTTCCAGGCTGTGGTGTCTTACTGACAGCATGGCTAGTTG 304
 DB 146 GGAATAATTCAGAGATATAGCGCGGAGACCTGTTGCATTTTTCACACAGCTTTGCTCG 205
 QY 305 GAGGTGCTAGCAACTGGAATTTCCAACTGCGCCCAATAGCTGGCTTCAATCTGGCAG 364
 DB 206 CATCAGCCACAACTGGATCTCCAACTGCGCCCAACACAGCTGGCTTCACTCTGGCGC 265
 QY 365 ATGCTGTTTGTAGCTGTGGATGGGAAACAGCAGGGAACCGCTGTCTCGAAAAACACA 424
 DB 266 ACGCGGGTAGAGCTGTGGCTGGGGAACACAGCGGGCAACACCTGGCGCAGGAGATC 325
 QY 425 AGACACTCTCCATAGACCAAGATGATTTCTGGGCTTTCAGTTATGATGATGGCTAGT 484
 DB 326 TGTAATCTGCGGCACTCCGTGCAATTTCTGGGCTTTCAGTTTGAAGATGGCTAAT 385
 QY 485 TTGACCTTCTGCAAGTAACTTTATTTTGGCAAAACGGGCGCAGGAAAGATCTATT 544
 DB 386 ATGACCTTCCGCGCACCACTGATTTTGAAGAAACGGGACAGGCAAGCTACAT 445

Search completed: February 19, 2004, 10:33:59
Job time : 84.5225 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2004 CompuGen Ltd.

OM nucleic - nucleic search, using sw model
Run on: February 19, 2004, 08:29:29 ; Search time 458.878 Seconds
(without alignments)
10186.874 Million cell updates/sec

Title: US-10-042-431-46
Perfect score: 1269
Sequence: 1 atgttggaacattgtcaag.....gacggtgagggcggtattg 1269

Scoring table: IDENTITY NUC
Gapop 10.0 , Gapext 1.0

Searched: 2449703 seqs, 1841816367 residues

Total number of hits satisfying chosen parameters: 4899406

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : Published Applications NA.*

- 1: /cgn2_6/prodata/1/pubna/US07_PUBCOMB.seq.*
- 2: /cgn2_6/prodata/1/pubna/US07_PUBCOMB.seq.*
- 3: /cgn2_6/prodata/1/pubna/US06_NEW_PUB.seq.*
- 4: /cgn2_6/prodata/1/pubna/US06_PUBCOMB.seq.*
- 5: /cgn2_6/prodata/1/pubna/US07_NEW_PUB.seq.*
- 6: /cgn2_6/prodata/1/pubna/US07_PUBCOMB.seq.*
- 7: /cgn2_6/prodata/1/pubna/US08_NEW_PUB.seq.*
- 8: /cgn2_6/prodata/1/pubna/US08_PUBCOMB.seq.*
- 9: /cgn2_6/prodata/1/pubna/US09A_PUBCOMB.seq.*
- 10: /cgn2_6/prodata/1/pubna/US09B_PUBCOMB.seq.*
- 11: /cgn2_6/prodata/1/pubna/US09C_PUBCOMB.seq.*
- 12: /cgn2_6/prodata/1/pubna/US09_NEW_PUB.seq.*
- 13: /cgn2_6/prodata/1/pubna/US10A_PUBCOMB.seq.*
- 14: /cgn2_6/prodata/1/pubna/US10B_PUBCOMB.seq.*
- 15: /cgn2_6/prodata/1/pubna/US10_NEW_PUB.seq.*
- 16: /cgn2_6/prodata/1/pubna/US10_NEW_PUB.seq.*
- 17: /cgn2_6/prodata/1/pubna/US60_NEW_PUB.seq.*
- 18: /cgn2_6/prodata/1/pubna/US60_PUBCOMB.seq.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	1269	100.0	1269	11	US-09-759-1308-416
2	1269	100.0	1269	14	US-10-042-431-46
3	1269	100.0	2044	11	US-09-759-1308-415
4	1269	100.0	2044	14	US-10-042-431-45
5	1267.4	99.9	2035	10	US-09-835-996A-20
6	1235.4	97.4	1384	10	US-09-835-996A-9
7	1122	88.4	1206	10	US-09-811-825-1
8	527.8	41.6	2481	10	US-09-880-107-3878
9	527.8	41.6	2858	15	US-10-198-846-13431
10	527.8	41.6	6952	13	US-10-133-013-144
11	525.4	41.4	2853	13	US-09-971-392-116
12	462.2	36.4	2506	12	US-10-108-260A-1262
13	456.8	36.0	1411	12	US-10-420-564-3
14	454.2	35.8	1194	13	US-10-312-088-7
15	453.6	35.7	1134	12	US-10-420-564-4

Sequence 3, Appli
Sequence 1, Appli
Sequence 1, Appli
Sequence 5, Appli
Sequence 9, Appli
Sequence 31, Appli
Sequence 1, Appli
Sequence 27, Appli
Sequence 29, Appli
Sequence 6, Appli
Sequence 9462, Ap
Sequence 29, Appli
Sequence 3, Appli
Sequence 4296, Ap
Sequence 6768, Ap
Sequence 4334, Ap
Sequence 4326, Ap
Sequence 4325, Ap
Sequence 1834, Ap
Sequence 1834, Ap
Sequence 253, App
Sequence 13032, A
Sequence 214253,
Sequence 214253,
Sequence 4162, Ap
Sequence 704, App
Sequence 3, Appli
Sequence 37615, A
Sequence 13508, A

ALIGNMENTS

RESULT 1

US-09-759-1308-416
; Sequence 416, Application US/097591308
; Publication No. US2003002279A1
; GENERAL INFORMATION:
; APPLICANT: Millennium Pharmaceuticals, Inc.
; APPLICANT: McCarthy, Sean A
; APPLICANT: Fraser, Christopher C
; APPLICANT: Sharp, John D
; APPLICANT: Barnes, Thomas S
; APPLICANT: Kirst, Susan J
; APPLICANT: Mackay, Charles R
; APPLICANT: Myers, Paul S
; APPLICANT: Leiby, Kevin R
; APPLICANT: Wrighton, Nicolas
; APPLICANT: Goodearl, Andrew
; APPLICANT: Holtzman, Douglas A
; TITLE OF INVENTION: NOVEL GENES ENCODING PROTEINS HAVING
; TITLE OF INVENTION: PROGNOSTIC, DIAGNOSTIC, PREVENTIVE, THERAPEUTIC, AND OTHER
; TITLE OF INVENTION: USES.
; FILE REFERENCE: MP100-5350MIM
; CURRENT APPLICATION NUMBER: US/09759,1308
; CURRENT FILING DATE: 2002-09-16
; PRIOR APPLICATION NUMBER: US 09/479,249
; PRIOR FILING DATE: 2000-01-07
; PRIOR APPLICATION NUMBER: US 09/559,497
; PRIOR FILING DATE: 2000-04-27
; PRIOR APPLICATION NUMBER: US 09/578,063
; PRIOR FILING DATE: 2000-05-24
; PRIOR APPLICATION NUMBER: US 09/333,159
; PRIOR FILING DATE: 1999-06-14
; PRIOR APPLICATION NUMBER: US 09/596,194
; PRIOR FILING DATE: 2000-07-14
; PRIOR APPLICATION NUMBER: US 09/342,364
; PRIOR FILING DATE: 1999-06-29
; PRIOR APPLICATION NUMBER: US 09/608,452
; PRIOR FILING DATE: 2000-06-30
; PRIOR APPLICATION NUMBER: US 09/393,996

```

; PRIOR FILING DATE: 1999-09-10
; PRIOR APPLICATION NUMBER: US 09/602,871
; PRIOR FILING DATE: 2000-06-23
; PRIOR APPLICATION NUMBER: US 09/420,707
; PRIOR FILING DATE: 1999-10-19
; NUMBER OF SEQ ID NOS: 460
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 416
; LENGTH: 1269
; TYPE: DNA
; ORGANISM: Homo sapiens
US-09-759-130B-416

Query Match      100.0%; Score 1269; DB 11; Length 1269;
Best Local Similarity 100.0%; Pred. No. 0;
Matches 1269; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 ATGTTGGAACCTTGTCCAGACAGTGGATTGTCTCCACACAGATGGAAATGTGGCTTCG 60
DB 1 ATGTTGGAACCTTGTCCAGACAGTGGATTGTCTCCACACAGATGGAAATGTGGCTTCG 60
QY 61 ATTCCTGGTGGGTATATGTTCCAGAGAAATGTGAATTCAGTACATATGCCAACTAAAGCT 120
DB 61 ATTCCTGGTGGGTATATGTTCCAGAGAAATGTGAATTCAGTACATATGCCAACTAAAGCT 120
QY 121 GTGACCCAGAGCAATTCATGAATATTAGTGAATATCATCCAAATCAAGGCTATCCCTGT 180
DB 121 GTGACCCAGAGCAATTCATGAATATTAGTGAATATCATCCAAATCAAGGCTATCCCTGT 180
QY 181 GAGGAATATGAGTGCACATGAGATGGGTATATCTTTCTGTTTACAGATTCCTCGA 240
DB 181 GAGGAATATGAGTGCACATGAGATGGGTATATCTTTCTGTTTACAGATTCCTCGA 240
QY 241 GGCCTAGTGCACCTTAAGACAGAGTTCCAGGCTCTGGGTCTTCTGACGATGGCCTA 300
DB 241 GGCCTAGTGCACCTTAAGACAGAGTTCCAGGCTCTGGGTCTTCTGACGATGGCCTA 300
QY 301 GTTGGAGTGTGACCACTGATTTCCAACTGCTGCCCAACATAGCTGGCTTCATCTG 360
DB 301 GTTGGAGTGTGACCACTGATTTCCAACTGCTGCCCAACATAGCTGGCTTCATCTG 360
QY 361 GCAGATGCTGTTTGGCTGTGATGGGGAACAGAGGGGAAACGGCTCTGCTCGAAAA 420
DB 361 GCAGATGCTGTTTGGCTGTGATGGGGAACAGAGGGGAAACGGCTCTGCTCGAAAA 420
QY 421 CACAAGACACTCTCCATAGACCAAGATGAGTTCTGGGCTTTTCACTTATGATGAGATGCT 480
DB 421 CACAAGACACTCTCCATAGACCAAGATGAGTTCTGGGCTTTTCACTTATGATGAGATGCT 480
QY 481 AGGTTTGACCTTCTCCAGTATGATAAATCTTATTTTGCAGAAAACGGGCCAGGAAAGATC 540
DB 481 AGGTTTGACCTTCTCCAGTATGATAAATCTTATTTTGCAGAAAACGGGCCAGGAAAGATC 540
QY 541 TATTATGTCGGCTATTCACAGGCAACCAACATGGGCTTTTATGCAATTTTCCACATGCCA 600
DB 541 TATTATGTCGGCTATTCACAGGCAACCAACATGGGCTTTTATGCAATTTTCCACATGCCA 600
QY 601 GAGCTGGCTCAGAAATCAAAATGATTTTCTGTTTGTGCTGCGAGATATGATGATGATGATGAT 660
DB 601 GAGCTGGCTCAGAAATCAAAATGATTTTCTGTTTGTGCTGCGAGATATGATGATGATGATGAT 660
QY 661 GCAAAAAGCCCGGACCAAAATTTTGTGCTGCGAGATATGATGATGATGATGATGATGATGAT 720
DB 661 GCAAAAAGCCCGGACCAAAATTTTGTGCTGCGAGATATGATGATGATGATGATGATGATGAT 720
QY 721 GCAAAAAGATTTTCTGATCAGACAGATTTCTCAGACAACTGTTTATTTTACCTTTGT 780
DB 721 GCAAAAAGATTTTCTGATCAGACAGATTTCTCAGACAACTGTTTATTTTACCTTTGT 780
QY 781 GGGCAGGTTGATTTCTGATCAGATTTGTAGTAATATCATGTTTCTGCGGTGGATTCACAC 840
DB 781 GGGCAGGTTGATTTCTGATCAGATTTGTAGTAATATCATGTTTCTGCGGTGGATTCACAC 840

```

RESULT 2

```

US-10-042-431-46
; Sequence 46, Application US/10042431
; Publication No. US20020182675A1
; GENERAL INFORMATION:
; APPLICANT: MCCARTHY, Sean A.
; APPLICANT: BARNES, Thomas M.
; APPLICANT: FRASER, Christopher C.
; APPLICANT: SHARP, John D.
; TITLE OF INVENTION: NOVEL GENES ENCODING PROTEINS HAVING DIAGNOSTIC,
; TITLE OF INVENTION: PREVENTIVE, THERAPEUTIC, AND OTHER USES
; FILE REFERENCE: 10147-6U2
; CURRENT APPLICATION NUMBER: US/10/042,431
; CURRENT FILING DATE: 2001-10-25
; PRIOR APPLICATION NUMBER: US 09/333,159
; PRIOR FILING DATE: 1999-06-14
; PRIOR APPLICATION NUMBER: US 09/578,063
; PRIOR FILING DATE: 2000-05-24
; NUMBER OF SEQ ID NOS: 79
; SOFTWARE: Patent In Ver. 2.1
; SEQ ID NO 46
; LENGTH: 1269
; TYPE: DNA
; ORGANISM: Homo sapiens
US-10-042-431-46

Query Match      100.0%; Score 1269; DB 14; Length 1269;
Best Local Similarity 100.0%; Pred. No. 0;
Matches 1269; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 ATGTTGGAACCTTGTCCAGACAGTGGATTGTCTCCACAGATGGAAATGTGGCTTCG 60
DB 1 ATGTTGGAACCTTGTCCAGACAGTGGATTGTCTCCACAGATGGAAATGTGGCTTCG 60
QY 61 ATTCCTGGTGGGTATATGTTCCAGAGAAATGTGAATTCAGTACATATGCCAACTAAAGCT 120
DB 61 ATTCCTGGTGGGTATATGTTCCAGAGAAATGTGAATTCAGTACATATGCCAACTAAAGCT 120
QY 121 GTGACCCAGAGCAATTCATGAATATTAGTGAATATCATCCAAATCAAGGCTATCCCTGT 180

```

Db 121 GTGACCCAGAGCATTTCATGATATTAGTGAATCATCCAAATCATCCAAATCAGGCTATCCCTGT 180
Qy 181 GAGGAATATGAGTGCAGCTGAAGATGGGTATATCCCTTTCGTGTAACAGAGATTCCTCGA 240
Db 181 GAGGAATATGAGTGCAGCTGAAGATGGGTATATCCCTTTCGTGTAACAGAGATTCCTCGA 240
Qy 241 GGCCTAGTGCACCTTAAGAGACAGAGTTCACAGGCTGTGGTGTACTGCAGCATGGGCTA 300
Db 241 GGCCTAGTGCACCTTAAGAGACAGAGTTCACAGGCTGTGGTGTACTGCAGCATGGGCTA 300
Qy 301 GTTGAGGTGCTAGCACTGAGTATCCAACTGCTCCCAATAGCCTGGGCTTCATCTG 360
Db 301 GTTGAGGTGCTAGCACTGAGTATCCAACTGCTCCCAATAGCCTGGGCTTCATCTG 360
Qy 361 CGAGATGCTGTTTGAAGTGTGGATGGGACACAGAGGGAAGCGCTGCTCGAAA 420
Db 361 CGAGATGCTGTTTGAAGTGTGGATGGGACACAGAGGGAAGCGCTGCTCGAAA 420
Qy 421 CACAAGACACTCTCCATAGACCAAGATGAGTTCCTGGGCTTCAGTATGATGAGATGGCT 480
Db 421 CACAAGACACTCTCCATAGACCAAGATGAGTTCCTGGGCTTCAGTATGATGAGATGGCT 480
Qy 481 AGTTTGAACCTTCCTGAGTGAATTAACCTTATTTGAGAAAACGGGCGAGAAAAGATC 540
Db 481 AGTTTGAACCTTCCTGAGTGAATTAACCTTATTTGAGAAAACGGGCGAGAAAAGATC 540
Qy 541 TATTATGCGGCTATTTCAGGGGACCAACCATGGGCTTTATTGCAATTTCCACCATGCCA 600
Db 541 TATTATGCGGCTATTTCAGGGGACCAACCATGGGCTTTATTGCAATTTCCACCATGCCA 600
Qy 601 GAGTGGCTCAGAAAATCAAAATGATTTTGTCTTTCAGCACCATAGCCACTGTTAAGCAT 660
Db 601 GAGTGGCTCAGAAAATCAAAATGATTTTGTCTTTCAGCACCATAGCCACTGTTAAGCAT 660
Qy 661 GCAAAAGCCCGGACCAAAATTTTGTCTGCGAGATATGATGATCAAGGATTTT 720
Db 661 GCAAAAGCCCGGACCAAAATTTTGTCTGCGAGATATGATGATCAAGGATTTT 720
Qy 721 GCAAAAAGAAATTTCTGTATCAGACCAAGATTTCTCAGACAACTTTGTTATTACCTTGT 780
Db 721 GCAAAAAGAAATTTCTGTATCAGACCAAGATTTCTCAGACAACTTTGTTATTACCTTGT 780
Qy 781 GGCAGGTGATTTCTGATCAGATTTGTAGTATATATCATTTACTTCTGGTGGATTCAC 840
Db 781 GGCAGGTGATTTCTGATCAGATTTGTAGTATATATCATTTACTTCTGGTGGATTCAC 840
Qy 841 ACCAAATATGAAATGACATGACGCGACAGTGTATGCTGCCACACTCTTCTGCGAACA 900
Db 841 ACCAAATATGAAATGACATGACGCGACAGTGTATGCTGCCACACTCTTCTGCGAACA 900
Qy 901 TCTGTGAAAATATTCTACCTGAGGCGAGGAGTGAATTCGTGTGAATCCGGGCAATT 960
Db 901 TCTGTGAAAATATTCTACCTGAGGCGAGGAGTGAATTCGTGTGAATCCGGGCAATT 960
Qy 961 GACTGGGGAGTGAGACCAAAATCTGGAATTCGAATTCGAATTCGAATTCGAATTCGAATTC 1020
Db 961 GACTGGGGAGTGAGACCAAAATCTGGAATTCGAATTCGAATTCGAATTCGAATTCGAATTC 1020
Qy 1021 AGAGTCAGAGATATGACGCTCCCTACAGCAATGTGGACAGGAGTTCAGGATGCTTCA 1080
Db 1021 AGAGTCAGAGATATGACGCTCCCTACAGCAATGTGGACAGGAGTTCAGGATGCTTCA 1080
Qy 1081 AATCAGAGAGTGAATGCTGCTCTGAGGAGTGAATTCGAATTCGAATTCGAATTCGAATTC 1140
Db 1081 AATCAGAGAGTGAATGCTGCTCTGAGGAGTGAATTCGAATTCGAATTCGAATTCGAATTC 1140
Qy 1141 ATTCTGATGGGCTCAGTGAATTTCTGCGGCTTTGATGCTCTCCCTCAGCGTATGTAC 1200
Db 1141 ATTCTGATGGGCTCAGTGAATTTCTGCGGCTTTGATGCTCTCCCTCAGCGTATGTAC 1200
Qy 1201 AATGAAATCATCCTATGATGACGAGGAGGACCAACCTTTCCAGGAGCGGTGTAG 1260
Db 1201 AATGAAATCATCCTATGATGACGAGGAGGACCAACCTTTCCAGGAGCGGTGTAG 1260

Qy 1261 GCGTATTG 1269
Db 1261 GCGTATTG 1269

RESULT 3

US-09-759-130B-415
; Sequence 415, Application US/09759130B
; Publication No. US2003022279A1
; GENERAL INFORMATION:
; APPLICANT: Millennium Pharmaceuticals, Inc.
; APPLICANT: McCarthy, Sean A
; APPLICANT: Fraser, Christopher C
; APPLICANT: Sharp, John D
; APPLICANT: Barnes, Thomas S
; APPLICANT: KIRST, Susan J
; APPLICANT: Mackay, Charles R
; APPLICANT: Myers, Paul S
; APPLICANT: Leiby, Kevin R
; APPLICANT: Wrighton, Nicolas
; APPLICANT: Goodearl, Andrew
; APPLICANT: Holtzman, Douglas A
; TITLE OF INVENTION: NOVEL GENES ENCODING PROTEINS HAVING
; TITLE OF INVENTION: PROGNOSTIC, DIAGNOSTIC, PREVENTIVE, THERAPEUTIC, AND OTHER
; FILE OF INVENTION: US
; FILE REFERENCE: MPI00-5350NM
; CURRENT APPLICATION NUMBER: US/09/759,130B
; CURRENT FILING DATE: 2002-09-16
; PRIOR APPLICATION NUMBER: US 09/479,249
; PRIOR FILING DATE: 2000-01-07
; PRIOR APPLICATION NUMBER: US 09/559,497
; PRIOR FILING DATE: 2000-04-27
; PRIOR APPLICATION NUMBER: US 09/578,063
; PRIOR FILING DATE: 2000-05-24
; PRIOR APPLICATION NUMBER: US 09/333,159
; PRIOR FILING DATE: 1999-06-14
; PRIOR APPLICATION NUMBER: US 09/596,194
; PRIOR FILING DATE: 2000-07-14
; PRIOR APPLICATION NUMBER: US 09/342,364
; PRIOR FILING DATE: 1999-06-29
; PRIOR APPLICATION NUMBER: US 09/608,452
; PRIOR FILING DATE: 2000-06-30
; PRIOR APPLICATION NUMBER: US 09/393,996
; PRIOR FILING DATE: 1999-09-10
; PRIOR APPLICATION NUMBER: US 09/602,871
; PRIOR FILING DATE: 2000-06-23
; PRIOR APPLICATION NUMBER: US 09/420,707
; PRIOR FILING DATE: 1999-10-19
; NUMBER OF SEQ ID NOS: 460
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 415
; LENGTH: 2044
; TYPE: DNA
; ORGANISM: Homo sapiens
US-09-759-130B-415

Query March 100.0%; Score 1269; DB 11; Length 2044;
Best Local Similarity 100.0%; Pred. No. 0;
Matches 1269; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 ATGTTGGAACCTGTGCAAGACAGTGGATTCTCACAGAGATGGAAATGGAATGGGCTTCTG 60
Db 126 ATGTTGGAACCTGTGCAAGACAGTGGATTCTCACAGAGATGGAAATGGAATGGGCTTCTG 185
Qy 61 ATTCTGCTGGCGTATATGTTCCAGAGAAATGGAATTCAGTACATATGCCAACTAAAGCT 120
Db 186 ATTCTGCTGGCGTATATGTTCCAGAGAAATGGAATTCAGTACATATGCCAACTAAAGCT 245
Qy 121 GTGGACCCAGAGACATTCATGAATATTAGTAAATCATCCAAATCATCAAGGCTATCCCTGT 180
Db 246 GTGGACCCAGAGACATTCATGAATATTAGTAAATCATCCAAATCATCAAGGCTATCCCTGT 305

181 GAGGATATGAGTGGGCACTGAAGATGGGTATATCTCTTCTGTTAAACAGGATTCCTCGA 240
 306 GAGGATATGAGTGGGCACTGAAGATGGGTATATCTCTTCTGTTAAACAGGATTCCTCGA 365
 241 GGCTAGTGCACCTAAGAGAGACAGGTTCCAGGCTGTGGTGTACTGAGCATGGCCTA 300
 366 GGCTAGTGCACCTAAGAGAGACAGGTTCCAGGCTGTGGTGTACTGAGCATGGCCTA 425
 301 GTTGGAGTGGTACGAACTGGATTTCCAACTGCCCAACAATAGCCTGGGCTTCATTCTG 360
 426 GTTGGAGTGGTACGAACTGGATTTCCAACTGCCCAACAATAGCCTGGGCTTCATTCTG 485
 361 GCAGATGCTGTTTGGAGTGGTGAAGGGGAAACAGAGGGGAAACGCTGTCTCGAAA 420
 486 GCAGATGCTGTTTGGAGTGGTGAAGGGGAAACAGAGGGGAAACGCTGTCTCGAAA 545
 421 CACAAGACACTCTCCATAGACCAAGATGAGTTCTGGGCTTTTCAGTTATGATGAGTGGT 480
 546 CACAAGACACTCTCCATAGACCAAGATGAGTTCTGGGCTTTTCAGTTATGATGAGTGGT 605
 481 AGTTTTCACCTCTCTGAGTGAATAAATTTATTTTCAGAAACAGAGGGGAAACGCTGTCTCGAAA 540
 606 AGTTTTCACCTCTCTGAGTGAATAAATTTATTTTCAGAAACAGAGGGGAAACGCTGTCTCGAAA 665
 541 TATTATGCTGGCTATTCCAGGGGCAACCATGGGCTTTTATTTGCAATTTTCCACCATGCCA 600
 666 TATTATGCTGGCTATTCCAGGGGCAACCATGGGCTTTTATTTGCAATTTTCCACCATGCCA 725
 601 GAGTGGCTCAGAAAATCAAAAATGATTTTTCCTTTAGCACCAATAGCCTGTTAGCAT 560
 726 GAGTGGCTCAGAAAATCAAAAATGATTTTTCCTTTAGCACCAATAGCCTGTTAGCAT 785
 661 GCAAAAAGCCCGGACCAAAATTTTCTGTCGACATATGATGATCAAGGATTTGTTT 720
 786 GCAAAAAGCCCGGACCAAAATTTTCTGTCGACATATGATGATCAAGGATTTGTTT 845
 721 GCGAAAAGAAATTTCTGTATCAGACAGATTTCTCAGAACATGTTGTTATTTACCTTTGT 780
 846 GCGAAAAGAAATTTCTGTATCAGACAGATTTCTCAGAACATGTTGTTATTTACCTTTGT 905
 781 GCGCAGGTGATTTCTGTATCAGATTTGTAGTAATATCATGTTTCTCTGGTGGATTTCAAC 840
 906 GCGCAGGTGATTTCTGTATCAGATTTGTAGTAATATCATGTTTCTCTGGTGGATTTCAAC 965
 841 ACCAACATATGAACATGAGCGGAGCAAGTGTATGCTGCCCACTCTCTGCTGGAACA 900
 966 ACCAACATATGAACATGAGCGGAGCAAGTGTATGCTGCCCACTCTCTGCTGGAACA 1025
 901 TCTGTGCAAAATATTTCTACATGAGCGGAGCAAGTGTATGCTGCCCACTCTCTGCTGGAACA 960
 1026 TCTGTGCAAAATATTTCTACATGAGCGGAGCAAGTGTATGCTGCCCACTCTCTGCTGGAACA 1085
 961 GACTGGGGGAGTGAGACCAAAAATCTGGAAAATGCAATCAGCCAACTCTCTGTAAGGTAC 1020
 1086 GACTGGGGGAGTGAGACCAAAAATCTGGAAAATGCAATCAGCCAACTCTCTGTAAGGTAC 1145
 1021 AGAGTCAGAGATATGACGGTCCCTACAGCAATGTGGACAGAGGTGAGGCTGAGGCTTTCA 1080
 1146 AGAGTCAGAGATATGACGGTCCCTACAGCAATGTGGACAGAGGTGAGGCTGAGGCTTTCA 1205
 1081 AATCCAGAGAGAGTGAAGTGTCTCTCTGAGGTGACCAACCTCTCTACCAATAAGAAAT 1140
 1206 AATCCAGAGAGAGTGAAGTGTCTCTCTGAGGTGACCAACCTCTCTACCAATAAGAAAT 1265
 1141 ATTCCTGAATGGGCTCAGTGGATTTTCACTGGGGTTTGGATGCTCTCAACCGTATGTAC 1200
 1266 ATTCCTGAATGGGCTCAGTGGATTTTCACTGGGGTTTGGATGCTCTCAACCGTATGTAC 1325
 1201 AATGAATATCATCCATCTGACAGAGGAGGAGACCACTTTTCCAGGGGAGCGTGTGAG 1260
 1326 AATGAATATCATCCATCTGATGAGGAGGAGGAGACCACTTTTCCAGGGGAGCGTGTGAG 1385
 1261 GCCGTATTG 1269

1386 GCCGTATTG 1394
 RESULT 4
 US-10-042-431-45
 ; Sequence 45, Application US/10042431
 ; Publication No. US20020182675A1
 ; GENERAL INFORMATION:
 ; APPLICANT: MCCARTHY, Sean A
 ; APPLICANT: BARNES, Thomas M
 ; APPLICANT: FRASER, Christopher C
 ; APPLICANT: SHARP, John D
 ; TITLE OF INVENTION: NOVEL GENES ENCODING PROTEINS HAVING DIAGNOSTIC,
 ; TITLE OF INVENTION: PREVENTIVE, THERAPEUTIC, AND OTHER USES
 ; FILE REFERENCE: 10147-6U2
 ; CURRENT APPLICATION NUMBER: US/10/042,431
 ; CURRENT FILING DATE: 2001-10-25
 ; PRIOR APPLICATION NUMBER: US 09/333,159
 ; PRIOR FILING DATE: 1999-06-14
 ; PRIOR APPLICATION NUMBER: US 09/578,063
 ; PRIOR FILING DATE: 2000-05-24
 ; NUMBER OF SEQ ID NOS: 79
 ; SOFTWARE: Patent in Ver. 2.1
 ; SEQ ID NO 45
 ; TYPE: DNA
 ; LENGTH: 2044
 ; ORGANISM: Homo sapiens
 ; US-10-042-431-45
 Query Match 100.0%; Score 1269; DB 14; Length 2044;
 Best Local Similarity 100.0%; Pred. No. 0;
 Matches 1269; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 QY 1 ATGTGGGAAACCTTGTCAAGACAGTGGATTTCTCACAGAAATGGAATGGGCTTCTG 60
 DB 126 ATGTGGGAAACCTTGTCAAGACAGTGGATTTCTCACAGAAATGGAATGGGCTTCTG 185
 QY 61 ATTCGTGGGCTATATGTTCCAGAGAAATGTGAATTCAGTACATATGCGCAATTAAGCT 120
 DB 186 ATTCGTGGGCTATATGTTCCAGAGAAATGTGAATTCAGTACATATGCGCAATTAAGCT 245
 QY 121 GTGGACCCAGAGCATTCATGAATATTAGTGAATATCCAAATCAAGGCTATCCCTGT 180
 DB 246 GTGGACCCAGAGCATTCATGAATATTAGTGAATATCCAAATCAAGGCTATCCCTGT 305
 QY 181 GAGGAATATGAAGTGGCAACTGGAAGTGGGTATATCTTTCTGTTAAACAGGATTCCTCGA 240
 DB 306 GAGGAATATGAAGTGGCAACTGGAAGTGGGTATATCTTTCTGTTAAACAGGATTCCTCGA 365
 QY 241 GGCCTAGTGCACCTTAAGAGACAGGTTCCAGGCTGTGGTGTACTGCGAGCATGGCCTA 300
 DB 366 GGCCTAGTGCACCTTAAGAGACAGGTTCCAGGCTGTGGTGTACTGCGAGCATGGCCTA 425
 QY 301 GTTGAGGTGTCTAGCAACTGGATTTCCAACTGCCCAACAATAGCCTGGGCTTCATTCTG 360
 DB 426 GTTGAGGTGTCTAGCAACTGGATTTCCAACTGCCCAACAATAGCCTGGGCTTCATTCTG 485
 QY 361 CGAGATGCTGTTTGTGAGTGGATGGGAAACAGAGGGGAAACGCTGTCTCGAAAA 420
 DB 486 CGAGATGCTGTTTGTGAGTGGATGGGAAACAGAGGGGAAACGCTGTCTCGAAAA 545
 QY 421 CACAAGACACTCTCCATAGACCAAGATGAGTTCTGGGCTTTTCAGTTATGATGAGTGGCT 480
 DB 546 CACAAGACACTCTCCATAGACCAAGATGAGTTCTGGGCTTTTCAGTTATGATGAGTGGCT 605
 QY 481 AGTTTTCACCTCTCTGAGTGAATAAATTTTTCAGAAACAGAGGGGAAACGCTGTCTCGAAA 540
 DB 606 AGTTTTCACCTCTCTGAGTGAATAAATTTTTCAGAAACAGAGGGGAAACGCTGTCTCGAAA 665
 QY 541 TATTATGCTGGCTATTCCAGGGGCAACCATGGGCTTTTATTTGCAATTTTCCACCATGCCA 600
 DB 666 TATTATGCTGGCTATTCCAGGGGCAACCATGGGCTTTTATTTGCAATTTTCCACCATGCCA 725

601 GAGCTGGCTCAGAAATCAAAATGATTTGCTTTAGCAACCATAGCCACTGTTAAGCAT 660
 Db GAGCTGGCTCAGAAATCAAAATGATTTGCTTTAGCAACCATAGCCACTGTTAAGCAT 785
 661 GCAAAAAGCCCGGACCAAAATTTTGTGCTGCCAGATATGATCAAGGGATTTGTT 720
 Db GCAAAAAGCCCGGACCAAAATTTTGTGCTGCCAGATATGATCAAGGGATTTGTT 845
 721 GCAAAAAGATTTCTGATCAGACCAAGATTTCTCAGACAACTTTGTTTACTTTGT 780
 Db GCAAAAAGATTTCTGATCAGACCAAGATTTCTCAGACAACTTTGTTTACTTTGT 905
 781 GCGCAGGTGATTTCTGATCAGATTTGTTAGTATATCATGTTACTTCTGGTGGATTCAC 840
 Db GCGCAGGTGATTTCTGATCAGATTTGTTAGTATATCATGTTACTTCTGGTGGATTCAC 965
 841 ACCAACAATGAAATGAGCGGACCAAGATTTGTTGCTGCCAGATATGATCAAGGGATTTGTT 900
 Db ACCAACAATGAAATGAGCGGACCAAGATTTGTTGCTGCCAGATATGATCAAGGGATTTGTT 1025
 901 TCTGTGCAAAATATTCTACCTGAGCGGACCAAGATTTGTTGCTGCCAGATATGATCAAGGGATTTGTT 960
 Db TCTGTGCAAAATATTCTACCTGAGCGGACCAAGATTTGTTGCTGCCAGATATGATCAAGGGATTT 1085
 961 GACTGGGGAGTGAGACCAAAATCTGGAATAATGCAATCAGCCAACTCTCTGTAAAGTAC 1020
 Db GACTGGGGAGTGAGACCAAAATCTGGAATAATGCAATCAGCCAACTCTCTGTAAAGTAC 1145
 1021 AGAGTCAGAGATATGACCGTCCCTACGAAATGTTGAGCAGAGGTCAGGACTGCTTTCA 1080
 Db AGAGTCAGAGATATGACCGTCCCTACGAAATGTTGAGCAGAGGTCAGGACTGCTTTCA 1205
 1081 AATCAGAAAGAGCTGAAATGCTGCTCTCTGAGGTGACCAACCTCATCTACCAATAAGAT 1140
 Db AATCAGAAAGAGCTGAAATGCTGCTCTCTGAGGTGACCAACCTCATCTACCAATAAGAT 1265
 1141 ATTCTGAATGGGCTCAGTGGAATTTATCTGGGGTTTGGATGCTCTCTACCGTATGTAC 1200
 Db ATTCTGAATGGGCTCAGTGGAATTTATCTGGGGTTTGGATGCTCTCTACCGTATGTAC 1325
 1201 AATGAAATCATCATCTGATGAGGAGGAGACCAACCTTTCCAGGGACGGTGTGAG 1260
 Db AATGAAATCATCATCTGATGAGGAGGAGACCAACCTTTCCAGGGACGGTGTGAG 1385
 1261 GCGGTATTG 1269
 Db GCGGTATTG 1394

RESULT 5
 US-09-835-996A-20
 ; Sequence 20, Application US/09835996A
 ; Patent No. US20020142953A1
 ; GENERAL INFORMATION:
 ; APPLICANT: Ballinger, Dennis
 ; APPLICANT: Loeb, Debra
 ; APPLICANT: Montgomery, Julie
 ; APPLICANT: Tang, Y. Tom
 ; APPLICANT: Zhou, Ping
 ; APPLICANT: Goodrich, Ryle
 ; APPLICANT: Liu, Chenghua
 ; APPLICANT: Asundi, Vinod
 ; APPLICANT: Zhao, Qing
 ; APPLICANT: Wehrman, Tom
 ; APPLICANT: Drmanac, Radoje
 ; APPLICANT: Ren, Feiyan
 ; APPLICANT: Qian, Xiaohong
 ; APPLICANT: Wang, Dunrui
 ; TITLE OF INVENTION: MATERIALS AND METHODS RELATING TO LIPID METABOLISM
 ; FILE REFERENCE: 28110/35915A
 ; CURRENT APPLICATION NUMBER: US/09/835,996A
 ; CURRENT FILING DATE: 2001-04-16

; PRIOR APPLICATION NUMBER: US 60/197,137
 ; PRIOR FILING DATE: 2000-04-14
 ; PRIOR APPLICATION NUMBER: US 09/714,936
 ; PRIOR FILING DATE: 2000-11-17
 ; PRIOR APPLICATION NUMBER: US 09/667,298
 ; PRIOR FILING DATE: 2000-09-22
 ; PRIOR APPLICATION NUMBER: US 09/631,451
 ; PRIOR FILING DATE: 2000-08-03
 ; PRIOR APPLICATION NUMBER: US 09/598,042
 ; PRIOR FILING DATE: 2000-06-20
 ; NUMBER OF SEQ ID NOS: 45
 ; SOFTWARE: PatentIn version 3.0
 ; SEQ ID NO 20
 ; LENGTH: 2035
 ; TYPE: DNA
 ; ORGANISM: Homo sapiens
 ; FEATURE:
 ; NAME/KEY: CDS
 ; LOCATION: (189)..(1415)
 ; US-09-835-996A-20

Query Match 99.9%; Score 1267.4; DB 10; Length 2035;
 Best Local Similarity 99.9%; Pred. No. 0;
 Matches 1268; Conservative 0; Mismatches 1; Indels 0; Gaps 0;
 QY 1 ATCTTGGAAACCTTGTCAAGACAGTGGATTGTCTCACACAGAAATGGAATGTGGCTTCTG 60
 Db 147 ATCTTGGAAACCTTGTCAAGACAGTGGATTGTCTCACACAGAAATGGAATGTGGCTTCTG 206
 QY 61 ATTCTGGTGGCTATATGTTCCAGAGAAATGGAATTCAGTACATATGCCAACTAAAGCT 120
 Db 207 ATTCTGGTGGCTATATGTTCCAGAGAAATGGAATTCAGTACATATGCCAACTAAAGCT 266
 QY 121 GTGGACCCAGAAAGCATTCATGAATATTAGTGAATCATCAACATCAAGGCTATCCCTGT 180
 Db 267 GTGGACCCAGAAAGCATTCATGAATATTAGTGAATCATCAACATCAAGGCTATCCCTGT 326
 QY 181 GAGGAATGAAATGCGAACTGAAAGATGGTATATCTTTCTGTAAACAGGAATCTCTGA 240
 Db 327 GAGGAATGAAATGCGAACTGAAAGATGGTATATCTTTCTGTAAACAGGAATCTCTGA 386
 QY 241 GGCTTAGTCAACCTAAGAGACAGGTTCCAGGCTTGTGCTGTACTGTGAGCATGSCCTA 300
 Db 387 GGCTTAGTCAACCTAAGAGACAGGTTCCAGGCTTGTGCTGTACTGTGAGCATGSCCTA 446
 QY 301 GTTGGAGGTGCTAGCAACTGGATTTCCAACTGCCCAACAATAGCTGGCTTCAATCTG 360
 Db 447 GTTGGAGGTGCTAGCAACTGGATTTCCAACTGCCCAACAATAGCTGGCTTCAATCTG 506
 QY 361 GCAGATGCTGGTTTTCAGCTGTGGATGGGGAACAGGAGGAAACGCTGTCTCGAAA 420
 Db 507 GCAGATGCTGGTTTTCAGCTGTGGATGGGGAACAGGAGGAAACGCTGTCTCGAAA 566
 QY 421 CACAGACACTCTCCATAGACCAAGATGATTTCTGGGCTTTTCAGTTATGATGAGATGGCT 480
 Db 567 CACAGACACTCTCCATAGACCAAGATGATTTCTGGGCTTTTCAGTTATGATGAGATGGCT 626
 QY 481 AGTTTGAACCTCTCTGAGTAAACTTTATTTTGCAGAAAACGGGGCAGGAAAAGATC 540
 Db 627 AGTTTGAACCTCTCTGAGTAAACTTTATTTTGCAGAAAACGGGGCAGGAAAAGATC 686
 QY 541 TATTATGTCGGCTATTTCAGAGGCAACCATGGGCTTTATTTGCAATTTTCCACCATGCCA 600
 Db 687 TATTATGTCGGCTATTTCAGAGGCAACCATGGGCTTTATTTGCAATTTTCCACCATGCCA 746
 QY 601 GAGCTGGCTCAGAAAATCAAAATGTTATTTGCTTTAGCACCATAGCCACTGTTAAGCAT 660
 Db 747 GAGCTGGCTCAGAAAATCAAAATGTTATTTGCTTTAGCACCATAGCCACTGTTAAGCAT 806
 QY 661 GCAAAAAGCCCGGACCAAAATTTTGTGCTGCCAGATATGATCAAGGGATTTGTT 720
 Db 807 GCAAAAAGCCCGGACCAAAATTTTGTGCTGCCAGATATGATCAAGGGATTTGTT 866

```

QY 721 GCGAAGAAATTTCTGTATCAGACAGATTTCTCAGACAACTTGTATTTACCTTTGT 780
Db 867 GCGAAGAAATTTCTGTATCAGACAGATTTCTCAGACAACTTGTATTTACCTTTGT 926
QY 781 GCGCAGGTGATTTCTGTATCAGATTTGTAGTAATATCATGTACTTCTGGGTGATTTCAAC 840
Db 927 GCGCAGGTGATTTCTGTATCAGATTTGTAGTAATATCATGTACTTCTGGGTGATTTCAAC 986
QY 841 ACCAACAATATGACATGAGCCGAGCAGTGTATATGCTGCCACACTTCTGCTGGAACA 900
Db 987 ACCAACAATATGACATGAGCCGAGCAGTGTATATGCTGCCACACTTCTGCTGGAACA 1046
QY 901 TCTGTGCAAAATATTTACACTGAGCCGAGCAGTGTATATGCTGCCACACTTCTGCTGGAACA 960
Db 1047 TCTGTGCAAAATATTTACACTGAGCCGAGCAGTGTATATGCTGCCACACTTCTGCTGGAACA 1106
QY 961 GACTGGGGAGTGAGACCAAAATCTGTGAAAAATGCAATCAGCAACTCTCTGTAGGTAC 1020
Db 1107 GACTGGGGAGTGAGACCAAAATCTGTGAAAAATGCAATCAGCAACTCTCTGTAGGTAC 1166
QY 1021 AGAGTCAGAGATATGAGCGTCCCTACAGCAATGTGGACAGAGGTGAGCACTGGCTTTCA 1080
Db 1167 AGAGTCAGAGATATGAGCGTCCCTACAGCAATGTGGACAGAGGTGAGCACTGGCTTTCA 1226
QY 1081 AATCCAGAGAGCGTGAATATGCTCTCTGTAGGTGACCAACTCTATACCAATAGAAAT 1140
Db 1227 AATCCAGAGAGCGTGAATATGCTCTCTGTAGGTGACCAACTCTATACCAATAGAAAT 1286
QY 1141 ATTCTCAATGGGTGAGCGTGGATTCATCTGGGTGAGTGTGATCTCTCAGCGTGTATG 1200
Db 1287 ATTCTCAATGGGTGAGCGTGGATTCATCTGGGTGAGTGTGATCTCTCAGCGTGTATG 1346
QY 1201 AATGAATATCATCTGATGAGCAGGAGGAGGAGCAACCTTTCCAGGAGCAGCGTGTGAG 1260
Db 1347 AATGAATATCATCTGATGAGCAGGAGGAGGAGCAACCTTTCCAGGAGCAGCGTGTGAG 1406
QY 1261 GCGGTATTTG 1269
Db 1407 GCGGTATTTG 1415

```

RESULT 6

```

; Sequence 9, Application US/09835996A
; Patent No. US20020142953A1
; GENERAL INFORMATION:
; APPLICANT: Ballinger, Dennis
; APPLICANT: Loeb, Debra
; APPLICANT: Montgomery, Julie
; APPLICANT: Tang, Y. Tom
; APPLICANT: Zhou, Ping
; APPLICANT: Goodrich, Ryle
; APPLICANT: Liu, Chenghua
; APPLICANT: Asundi, Vinod
; APPLICANT: Zhao, Qing
; APPLICANT: Wehrman, Tom
; APPLICANT: Drmanac, Radoje
; APPLICANT: Ren, Feiyan
; APPLICANT: Qian, Xiahong
; APPLICANT: Wang, Dunrui
; TITLE OF INVENTION: MATERIALS AND METHODS RELATING TO LIPID METABOLISM
; FILE REFERENCE: 28110/35915A
; CURRENT APPLICATION NUMBER: US/09/835,996A
; PRIOR FILING DATE: 2001-04-15
; PRIOR APPLICATION NUMBER: US 60/197,137
; PRIOR FILING DATE: 2000-04-14
; PRIOR APPLICATION NUMBER: US 09/714,936
; PRIOR FILING DATE: 2000-11-17
; PRIOR APPLICATION NUMBER: US 09/667,298
; PRIOR FILING DATE: 2000-09-22
; PRIOR APPLICATION NUMBER: US 09/631,451
; PRIOR FILING DATE: 2000-08-03
; PRIOR APPLICATION NUMBER: US 09/598,042

```

```

; PRIOR FILING DATE: 2000-06-20
; NUMBER OF SEQ ID NOS: 45
; SOFTWARE: Patent in version 3.0
; SEQ ID NO 9
; LENGTH: 1384
; TYPE: DNA
; ORGANISM: Homo sapiens
; US-09-835-996A-9

Query Match 97.4%; Score 1235.4; DB 10; Length 1384;
Best Local Similarity 99.9%; Pred. No. 0;
Matches 1236; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 ATGTTGGAACCTTGTCAAGACAGTGTGATTTCTCAGACAGATTTGTCTCAGACAGATTTGGCTTCTG 60
Db 147 ATGTTGGAACCTTGTCAAGACAGTGTGATTTGTCTCAGACAGATTTGGCTTCTG 206
QY 61 ATTCGTGTGGGTATATGTTTCCAGAGAAATGTGAATTCAGTACATATGCCAATTAAGCT 120
Db 207 ATTCGTGTGGGTATATGTTTCCAGAGAAATGTGAATTCAGTACATATGCCAATTAAGCT 266
QY 121 CTGGACCCAGAGCAATTCATGATATTAGTGAATCATCCAAATCAATCAAGGCTATCCCTGT 180
Db 267 GTGGACCCAGAGCAATTCATGATATTAGTGAATCATCCAAATCAATCAAGGCTATCCCTGT 326
QY 181 GAGGAATATGAAGTCGCAACTGAAGATGGGTATATCCTTTCTGTAAACAGGATTCCTCGA 240
Db 327 GAGGAATATGAAGTCGCAACTGAAGATGGGTATATCCTTTCTGTAAACAGGATTCCTCGA 386
QY 241 GGCCTAGTGCACCTAAGAGACAGAGTTCAGAGGCTGTGGTGTACTGACAGCATGGCCTA 300
Db 387 GGCCTAGTGCACCTAAGAGACAGAGTTCAGAGGCTGTGGTGTACTGACAGCATGGCCTA 446
QY 301 GTTGGAGGTGCTAGCAACTGGATTTCCAACTGCCCAACAAATAGCTTGGCTTCAATTCG 360
Db 447 GTTGGAGGTGCTAGCAACTGGATTTCCAACTGCCCAACAAATAGCTTGGCTTCAATTCG 506
QY 361 GCAGATGCTGTTTGTAGCTGTGATGGGAAACAGCAGGGGAAACGGCTGGTCTGAAAA 420
Db 507 GCAGATGCTGTTTGTAGCTGTGATGGGAAACAGCAGGGGAAACGGCTGGTCTGAAAA 566
QY 421 CACAGACACTCTCCATAGACCAAGATGAGTTCTGGGCTTTCAGTTATGATGAGATGGCT 480
Db 567 CACAGACACTCTCCATAGACCAAGATGAGTTCTGGGCTTTCAGTTATGATGAGATGGCT 626
QY 481 AGGTTTGACCTTCTGTGAGTGAATAAATTTATTTTGCAGAAAAACGGGCGCAGAAAAAGATC 540
Db 627 AGGTTTGACCTTCTGTGAGTGAATAAATTTATTTTGCAGAAAAACGGGCGCAGAAAAAGATC 686
QY 541 TATTATGTCGGCTATTTCACAGGCAACCAACCATGGCTTTTATGCAATTTCCACCATGCCA 600
Db 687 TATTATGTCGGCTATTTCACAGGCAACCAACCATGGCTTTTATGCAATTTCCACCATGCCA 746
QY 601 GAGCTGGCTCAGAAAAATCAAAATGTATTTTGTAGTCCCAATAGCCATGTTTAAAGCAT 660
Db 747 GAGCTGGCTCAGAAAAATCAAAATGTATTTTGTAGTCCCAATAGCCATGTTTAAAGCAT 806
QY 661 GCAAAAAGCCCGGAGCAAAATTTTGTGTCGAGATATGATGATCAAGGATTTGTT 720
Db 807 GCAAAAAGCCCGGAGCAAAATTTTGTGTCGAGATATGATGATCAAGGATTTGTT 866
QY 721 GCGAAAAAGAAATTTCTGTATCAGACCAAGATTTCTCAGACAACTTGTATTTACCTTTGT 780
Db 867 GCGAAAAAGAAATTTCTGTATCAGACCAAGATTTCTCAGACAACTTGTATTTACCTTTGT 926
QY 781 GCGCAGGTGATTTCTGTATCAGATTTGTAGTAATATCATGTACTTCTGGGTGATTTCAAC 840
Db 927 GCGCAGGTGATTTCTGTATCAGATTTGTAGTAATATCATGTACTTCTGGGTGATTTCAAC 986
QY 841 ACCAACAATATGACATGAGCCGAGCAGTGTATATGCTGCCACACTTCTGCTGGAACA 900
Db 987 ACCAACAATATGACATGAGCCGAGCAGTGTATATGCTGCCACACTTCTGCTGGAACA 1046

```

QY 901 TCTGTCAAAATATCTCACTGGAGCCAGGAGTGAATCTTGTGTGAATCTCGGCAATTT 960
 DB 1047 TCTGTCAAAATATCTCACTGGAGCCAGGAGTGAATCTTGTGTGAATCTCGGCAATTT 1106
 QY 961 GACTGGGGAGTGAGACCAAAATCTGGAATAATCAATCAACCACTCTGTAAAGTAC 1020
 DB 1107 GACTGGGGAGTGAGACCAAAATCTGGAATAATCAATCAACCACTCTGTAAAGTAC 1166
 QY 1021 AGAGTCAGAGATATGAGCTCCCTACAGCAATGTGACAGGAGGTCAAGACTGGCTTTCA 1080
 DB 1167 AGAGTCAGAGATATGAGCTCCCTACAGCAATGTGACAGGAGGTCAAGACTGGCTTTCA 1226
 QY 1081 AATCCAGAGAGCTGAAAATGTCTCTCTGAGGTGACCAACCTCATCTACCATTAAGAAAT 1140
 DB 1227 AATCCAGAGAGCTGAAAATGTCTCTCTGAGGTGACCAACCTCATCTACCATTAAGAAAT 1286
 QY 1141 ATTCCTGAATGGGCTCACTGAGATTTTCATCTGGGTTTGGATGCTCTCACCCTATGATAC 1200
 DB 1287 ATTCCTGAATGGGCTCACTGAGATTTTCATCTGGGTTTGGATGCTCTCACCCTATGATAC 1346
 QY 1201 AATGAATCATCTCATCTGATGACAGGAGGAGACCA 1237
 DB 1347 AATGAATCATCTCATCTGATGACAGGAGGAGACCA 1383

RESULT 7

US-09-811-825-1
 ; Sequence 1, Application US/09811825
 ; Patent No. US2002014297A1
 ; GENERAL INFORMATION:
 ; APPLICANT: YAN, Chunhua et al.
 ; TITLE OF INVENTION: ISOLATED HUMAN LIPASE PROTEINS, NUCLEIC
 ; TITLE OF INVENTION: ACID MOLECULES ENCODING HUMAN LIPASE PROTEINS, AND USES
 ; FILE REFERENCE: THEREOF
 ; FILE REFERENCE: CL001170
 ; CURRENT APPLICATION NUMBER: US/09/811,825
 ; CURRENT FILING DATE: 2001-03-20
 ; NUMBER OF SEQ ID NOS: 4
 ; SOFTWARE: FastSeq for Windows Version 4.0
 ; SEQ ID NO 1
 ; LENGTH: 1206
 ; TYPE: DNA
 ; ORGANISM: Human
 ; US-09-811-825-1

Query Match 88.4%; Score 1122; DB 10; Length 1206;

Best Local Similarity 100.0%; Pred. No. 0;

Matches 1122; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 148 AGTGAATCATCAACATCAAGCTATCCCTGTGAGGATATCAAGTCCCACTGAAGAT 207
 DB 76 AGTGAATCATCAACATCAAGCTATCCCTGTGAGGATATCAAGTCCCACTGAAGAT 135
 QY 208 GGGTATATCTTCTGTAAACAGGATCTCTCGAGGCTTAGTGAACCAAGACAGGT 267
 DB 136 GGGTATATCTTCTGTAAACAGGATCTCTCGAGGCTTAGTGAACCAAGACAGGT 195
 QY 268 TCAGGCTGTGTGTATCTGACAGATGGCTTAGTTGGAGGTCTAGCAACTGGATTTCC 327
 DB 196 TCAGGCTGTGTGTATCTGACAGATGGCTTAGTTGGAGGTCTAGCAACTGGATTTCC 255
 QY 328 AACCTGCCAACAAATAGCTGGCTTCATCTGGCAGATGCTGGTTTTCACCTGTGGATG 387
 DB 256 AACCTGCCAACAAATAGCTGGCTTCATCTGGCAGATGCTGGTTTTCACCTGTGGATG 315
 QY 388 GGGAAACAGCAGGGGAAACGCTGGTCTCGAAAACCAAGACACTCTCCATAGACCAAGAT 447
 DB 316 GGGAAACAGCAGGGGAAACGCTGGTCTCGAAAACCAAGACACTCTCCATAGACCAAGAT 375
 QY 448 GAGTTCTGGGCTTTCAGTTATGATGATGGCTTAGTTGACCTTCTCCTCAGTGAATAC 507
 DB 376 GAGTTCTGGGCTTTCAGTTATGATGATGGCTTAGTTGACCTTCTCCTCAGTGAATAC 435

QY 508 TTTATTTTTCAGAAAACGGCCAGGAAAGATCTATTATGTGGCTATTTCACAGGCGACC 567
 DB 436 TTTATTTTTCAGAAAACGGCCAGGAAAGATCTATTATGTGGCTATTTCACAGGCGACC 495
 QY 568 ACCATGGGCTTTATTTGCATTTTCCCATGCGAGAGCTGGCTCAGAAAATCAAAATGTAT 627
 DB 496 ACCATGGGCTTTATTTGCATTTTCCCATGCGAGAGCTGGCTCAGAAAATCAAAATGTAT 555
 QY 628 TTTGCTTTTACACCCATAGCACTGTTTAGCATGCAAAAAGCCCGGACCAATTTTTCG 687
 DB 556 TTTGCTTTTACACCCATAGCACTGTTTAGCATGCAAAAAGCCCGGACCAATTTTTCG 615
 QY 688 TTGCTGCCAGATATGATGATCAAGGATTTCTTTGGCAAAAAGAAATTTCTGTATCAGACC 747
 DB 616 TTGCTGCCAGATATGATGATCAAGGATTTCTTTGGCAAAAAGAAATTTCTGTATCAGACC 675
 QY 748 AGATTCTTCAGACAACTTTGTTATTTTACCTTTGTGGCCAGGTGATTTCTGTATCAGATTGT 807
 DB 676 AGATTCTTCAGACAACTTTGTTATTTTACCTTTGTGGCCAGGTGATTTCTGTATCAGATTGT 735
 QY 808 AGTAATATCATGTTTACTTTCTGGGTGATTCACACCAACAATATGAACATGAGCCGAGCA 867
 DB 736 AGTAATATCATGTTTACTTTCTGGGTGATTCACACCAACAATATGAACATGAGCCGAGCA 795
 QY 868 AGTGTATATGCTGCCACACTCTTTGCTGGAAACATCTGTGCAAAATATTTTACACTGGAGC 927
 DB 796 AGTGTATATGCTGCCACACTCTTTGCTGGAAACATCTGTGCAAAATATTTTACACTGGAGC 855
 QY 928 CAGGCACTGATTTCTGCTGAACCTCCGGGCAATTTGACTGGGGAGTGAGACCAAAATCTG 987
 DB 856 CAGGCACTGATTTCTGCTGAACCTCCGGGCAATTTGACTGGGGAGTGAGACCAAAATCTG 915
 QY 988 GAAAAATGCAATCAGCCAACTCTCTGTAAGGTACAGAGTACAGAGATATGACGGTCCCTTACA 1047
 DB 916 GAAAAATGCAATCAGCCAACTCTCTGTAAGGTACAGAGTACAGAGATATGACGGTCCCTTACA 975
 QY 1048 GCAATGTGGACAGAGGTGAGACTGGCTTTTCAATCCAGAGAGCTGAAAATGTGCTC 1107
 DB 976 GCAATGTGGACAGAGGTGAGACTGGCTTTTCAATCCAGAGAGCTGAAAATGTGCTC 1035
 QY 1108 TCTGAGGTGACCAACCTCATCTACCAATAAGAAATTTCTGAAATGGGCTCACGTTGGATTTC 1167
 DB 1036 TCTGAGGTGACCAACCTCATCTACCAATAAGAAATTTCTGAAATGGGCTCACGTTGGATTTC 1095
 QY 1168 ATCTGGGGTTTGGATGCTCTCACCCTATGTAATAAGAAATTCATCTGATGCGAGCAG 1227
 DB 1096 ATCTGGGGTTTGGATGCTCTCACCCTATGTAATAAGAAATTCATCTGATGCGAGCAG 1155
 QY 1228 GAGGAGACCAACCTTTTCCAGGAGCGGTGTGAGGCGGTATTG 1269
 DB 1156 GAGGAGACCAACCTTTTCCAGGAGCGGTGTGAGGCGGTATTG 1197

RESULT 8

US-09-880-107-3878
 ; Sequence 3878, Application US/09880107
 ; Patent No. US20020142981A1
 ; GENERAL INFORMATION:
 ; APPLICANT: Horne, Darci T.
 ; APPLICANT: Vockley, Joseph G.
 ; APPLICANT: Scherf, Uwe
 ; APPLICANT: Gene Logic, Inc.
 ; TITLE OF INVENTION: Gene Expression Profiles in Liver Cancer
 ; CURRENT APPLICATION NUMBER: US/09/880,107
 ; CURRENT FILING DATE: 2001-06-14
 ; PRIOR APPLICATION NUMBER: US 60/211,379
 ; PRIOR FILING DATE: 2000-06-14
 ; PRIOR APPLICATION NUMBER: US 60/237,054
 ; PRIOR FILING DATE: 2000-10-02
 ; NUMBER OF SEQ ID NOS: 3950
 ; SOFTWARE: Patent In Ver. 2.1
 ; SEQ ID NO 3878

```

; LENGTH: 2481
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; OTHER INFORMATION: Genbank Accession No. US20020142981A1 Z31690
US-09-880-107-3878

```

```

Query Match 41.6%; Score 527.8; DB 10; Length 2481;
Best Local Similarity 66.6%; Pred. No. 4.9e-162;
Matches 751; Conservative 0; Mismatches 376; Indels 0; Gaps 0;

```

QY	115	AAAGCTGTGAGCCAGAGCAATTCATGAATATTAGTGAATATCAATCAATCAAGGCTAT	174
DB	112	ACAGCTGTGATCTCTGAAACAAACATGAATGATGAATATCTCTTACTGCGGATTC	171
QY	175	CCCTGTGAGGAATATGAGTCGCACTGAAGATGGGTATATCTTCTGTAAACAGATT	234
DB	172	CCTAGTGAAGGAATACCTAGTTGAGACAGAGATGATATATCTGTGCCCTTAAACCAAT	231
QY	235	CCTGAGGCTAGTGCAACCTTAAGAGACAGGTTCCAGGCTGTGGTGTACTGACAGCAT	294
DB	232	CCTCATGGAGGAGAACCACTCTGACAAAGTCCAAACAGTTGTCTTCTGCAACAT	291
QY	295	GGCTAGTGTGAGGTCTAGCACTGATTTCCAACTGCGCCCAACAAATAGCTGGGCTTC	354
DB	292	GGCTTGTGCGAGATTTCTAGTAATCTGGCTCAAAACCTTGCCAAACAGCAGCTGGGCTTC	351
QY	355	ATTCCTGGCAGATGCTGGTGTGAGTGTGGATGGGAAACAGCAGGGAACGCTGGTCT	414
DB	352	ATTCCTGTGATGCTGGTGTGAGTGTGGATGGGAAACAGCAGGGAATACCTGGTCT	411
QY	415	CGAAACACAGACACTCTCCATAGACCAAGATGAGTTCTGGGCTTTTCAATGATGAG	474
DB	412	CGAAACACATAGCACTCTCAGTTCTCAGGATGAATCTGGGCTTTTCAATGATGAG	471
QY	475	ATGGCTAGTTTGACCTTCTGCACTGATGAATCTTTATTTGACAGAAACGGGCGAGAA	534
DB	472	ATGGCAAAATAGCACTTACCACTGATGAATCTTTATTTGACAGAAACGGGCGAGAA	531
QY	535	AAGATCTATTATGTGGCTATTACAGGCGCAACCACTGGCTTTATTTGATTTTCCACC	594
DB	532	CAAGTGATATTATGTGGCTATTCTCAAGGCGCAACCACTAGATTTTATGATTTTACAG	591
QY	595	ATGCCAGAGCTGGCTCAGAAATTCMAAATGATTTTGTCTTAGCAGCCATAGCCAGCTTT	654
DB	592	ATCCCTGAGCTGGCTTAAAGGATTAATATGTTTTTGGCTGGGCTCTGGCTTCCGTC	651
QY	655	AAGCATGCAAAAGCCCGGACCAAAATTTTGTGCTGCGCAGATATGATGATCAAGGGA	714
DB	652	GCCTTCTGTACTAGCCCTATGGCCAAATTAGGACGATTTACAGATCATCTCTTAAGAC	711
QY	715	TTGTTTGGCAAAAGAAATTTCTGATCAGACCAAGATTTCTCAGACAACTTTGTTATTAC	774
DB	712	TTATTGGAGACAAAGAAATTTCTCCAGAGTGGCTTTTGAAGTGGCTGGGTACCCAC	771
QY	775	CTTTGGCCAGGTGATTTCTGATCAGATTTGTAGTATATCATGTTACTTCTGGGTGGA	834
DB	772	GTTTGCACTCATGCTATGAGGAGCTCTGTGGAATCTCTGTCTTCTGTGTGGA	831
QY	835	TTCAACACCAAAATATGAACATGAGCGGACAGTGTATATGCTGCCCACTCTTGCT	894
DB	832	TTTAATGAGAGAAATTTAAATATGTCTAGAGTGGATGTATATACACATCTCTCTGCT	891
QY	895	GGAACTCTGTGCAAAATTTCTACCTGGAGCCAGGAGTGAATTTCTGTGGAATCTCGG	954
DB	892	GGAACTCTGTGCAAAATTTCTACCTGGAGCCAGGCTGTAAATTTCAAAAGTTTCAA	951
QY	955	GCATTGTAGTGGGAGTGAGACCAAAATCTGGAATAATGCAATTCAGCAACTCTGTA	1014
DB	952	GCCTTTGACTGGGAGAGAGTCCCAAGAAATTTATTTTCAATACAGAGTTATCTCC	1011
QY	1015	AGGTACAGAGTCAGAGATATGACGCTCCCTACAGCAATGTGGACAGGAGGTGAGACTGG	1074

RESULT 9

```

US-10-198-846-13431
; Sequence 13431, Application US/10198846
; Publication No. US2003009974A1
; GENERAL INFORMATION:
; APPLICANT: Lillie, James
; APPLICANT: Xu, Yongyao
; APPLICANT: Wang, Youzhen
; APPLICANT: Steinmann, Kathleen
; TITLE OF INVENTION: NOVEL GENES, COMPOSITIONS, KITS, AND METHODS
; TITLE OF INVENTION: FOR IDENTIFICATION, ASSESSMENT, PREVENTION, AND
; TITLE OF INVENTION: THERAPY OF BREAST CANCER
; FILE REFERENCE: MRI-049
; CURRENT APPLICATION NUMBER: US/10/198,846
; PRIOR FILING DATE: 2002-07-18
; PRIOR APPLICATION NUMBER: 60/306,220
; PRIOR FILING DATE: 2001-07-18
; NUMBER OF SEQ ID NOS: 14084
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 13431
; LENGTH: 2858
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: misc feature
; LOCATION: 1, 2, 2718, 2720, 2721, 2722, 2723, 2846, 2847, 2848,
; LOCATION: 2849, 2851, 2852, 2853, 2854, 2855, 2856, 2857, 2858
; OTHER INFORMATION: n = A,T,C or G
US-10-198-846-13431

```

```

Query Match 41.6%; Score 527.8; DB 15; Length 2858;
Best Local Similarity 66.6%; Pred. No. 5.4e-162;
Matches 751; Conservative 0; Mismatches 376; Indels 0; Gaps 0;

```

QY	115	AAAGCTGTGAGCCAGAGCAATTCATGAATATTAGTGAATATCAATCAATCAAGGCTAT	174
DB	247	ACAGCTGTGATCTCTGAAACAAACATGAATGATGATGAATATCTCTTACTGGGATTC	306
QY	175	CCCTGTGAGGAATATGAAGTCCAACTGAAGATGGGTATATCTTCTTGTAAACAGGATT	234
DB	307	CCTAGTGAAGATACCTAGTTGAGACAGAGATGATATATCTGTGCTTTAACCGAATT	366
QY	235	CCTGAGGCTAGTGCAACCTTAAGAGACAGGTTCCAGGCTGTGGTGTACTGACGAT	294
DB	367	CCTCATGGGAGGAAGAACCAATCTTGACAAAGGTCCCAACCAAGTTGTCTTCTGCAACAT	426
QY	295	GGCTAGTTGGAGGTGCTAGCAACTGGATTTTCCAACTGCCCAACAAATAGCTGGGCTTC	354
DB	427	GGCTTCTGCGAGATTTCTAGTAACTGGGTCAAACTTGGCCAAACAGAGCTTGGGCTTC	486
QY	355	ATTCTGGCAGATGCTGGTTTTGACCGTGTGGATGGGAAACAGCAGGGGAAACGCTGGTCT	414
DB	487	ATTCTTGTCTGCTGGTGTGGTGTGGATGGGCAACAGCAGAGGAATAACCTGGTCT	546
QY	415	CGAAACACAGACACTCTCCATAGACCAAGATGATGTTCTGGGCTTTTCAATGATGAG	474
DB	547	CGAAACATAGACACTCTCAGTTCTTCTCAGGATGAATTTCTGGGCTTTTCAATGATGAG	606

QY 475 ATGGCTAGGTTTGACCTCTGCGAGTGATTAACCTTTTTCAGAGAAACGGCCAGAA 534
Db |||||
QY 607 ATGGCAAAATATGACCTACCAAGCTTCCATTAACCTTCTGTAATAAATCGGCGCAAGAA 666
Db |||||
QY 535 AAGATCTATTATGTCGGCTATTTCAGGGCCACCACTATGGCTTTTATGCAATTTCCACC 594
Db |||||
QY 667 CAAAGTGATATTATGCGGTCACTTCAAGGCCACCACTATAGGTTTATAGCATTTTCACAG 726
Db |||||
QY 595 ATGCCAGAGCTGGCTCAGAAAATCAAAATGATTTTGTCTTTAGCACCCTATAGCCATGCTTT 654
Db |||||
QY 727 ATCCCTGAGCTGGCTGAAAAGATTAATGTTTGTGCTGGCTGCTGCTGGCTTCCGTC 786
Db |||||
QY 655 AAGCATGCAAAAAGCCCGGACCAAAATTTTGTGCTGCCAGATATGATGATCAAGGGA 714
Db |||||
QY 787 GCCCTCTGCTAGCCCTATGCCCAATTAGGAGCATACAGATCATCTCATTAAGGAC 846
Db |||||
QY 715 TTGTTGGCAAAAAGAAATTTCTGTATGACCAAGATTTCTCAGCAAACTTTGTTATTAC 774
Db |||||
QY 847 TTAATTTGAGACAAAGAAATTTCTCCCAAGAGTGGCTTTTGAAGTGGCTGGGTACCCAC 906
Db |||||
QY 775 CTTTGTGCCAGGTGATTTCTGATCAGATTTGTAGTATATATCATGTTACTCTGGGTGA 834
Db |||||
QY 907 GTTTGCACTCATGCTACACTGAGGAGCTCTGTGAAATCTCTGTTTCTCTGTGGGA 966
Db |||||
QY 835 TTCAACACCAAAATATGAACATGAGCGGACCAAGTGTATATGCTGCCCACTCTTTGCT 894
Db |||||
QY 967 TTTAATGAGAGAAATTTAAATATGCTAGAGTGGATGTATATACACACATTTCTCTGCT 1026
Db |||||
QY 895 GGAACATCTGCAAAATATTTCTACCTGGAGCGAGCAGTGAATTTCTGTGTAACCTCGG 954
Db |||||
QY 1027 GGAATCTCTGTGCAAAATGTTTACCTGGAGCAGGCTGTAAATTTCCAAAAGTTTCAA 1086
Db |||||
QY 955 GCATTTGACTGGGGAGTGAGACCAAAATCTGGAAAATGCAATCAGCCAACTCTCTGTA 1014
Db |||||
QY 1087 GCCCTTTGACTGGGAAGCAGTGCACCAAGATTTATTTTCTATACACAGATTTCTCTCC 1146
Db |||||
QY 1015 AGGTACAGAGTCAGAGATATGACGGTCCCTACAGCAATGGGACAGAGTCAGGACTGG 1074
Db |||||
QY 1147 ACATACAAATGGAAGGACATCTTTGTGCGGACTGCTGAGCGGGGTGACGACTGG 1206
Db |||||
QY 1075 CTTTCAAATCAGAAAGACGTGAAATGCTGCTCTGAGGTGACCAACCTCATCTACAT 1134
Db |||||
QY 1207 CTTCAGAGTCTACGAGTCAATATCTTACTGACTCAGATCACCACATTTGGTTCAT 1266
Db |||||
QY 1135 AAGAAATATCTGAATGGCTCAGCGTGAATTTCACTGGGTTTGGATGCTCTCTCACCGT 1194
Db |||||
QY 1267 GAGAGCAATCCGGAATGGGAGCATCTTGACTTTCAATTTGGGCTGGATGCCCTTGGAG 1326
Db |||||
QY 1195 ATGTACATGAATCATCCATCTGATGCGAGGAGGAGGACCACT 1241
Db |||||
QY 1327 CTTTATAAATAATTTAATCTAATGAGGAATATCAGTGAAAGCT 1373
Db |||||

RESULT 10

US-10-133-013-144
; Sequence 144, Application US/10133013
; Publication No. US20030166903A1
; GENERAL INFORMATION:
; APPLICANT: Astromoff, Anna
; APPLICANT: Bandman, Olga
; APPLICANT: Cocks, Benjamin G.
; TITLE OF INVENTION: GENES ASSOCIATED WITH VASCULAR DISEASE
; CURRENT APPLICATION NUMBER: US/10/133,013
; CURRENT FILING DATE: 2002-04-25
; PRIOR APPLICATION NUMBER: 60/287,067
; PRIOR FILING DATE: 2001-04-27
; NUMBER OF SEQ ID NOS: 271
; SOFTWARE: PERL Program
; SEQ ID NO 144
; LENGTH: 6952
; TYPE: DNA
; ORGANISM: Homo sapiens

; FEATURE:

; NAME/KEY: misc_feature
; OTHER INFORMATION: Incyte ID No. US20030166903A1 410917.13
US-10-133-013-144

Query Match 41.6%; Score 527.8; DB 13; Length 6952;
Best Local Similarity 66.6%; Pred No. 9.9e-162; Indels 0; Gaps 0;
Matches 751; Conservative 0; Mismatches 376;
QY 115 AAAGCTGTGCGACCCAGAGCAATTCATGAAATATTAGTGAATCATCAACATCAAGGCTAT 174
Db 4577 ACAGCTGTGATCTCTGAAACAAACATGAATGTAGTGAATATTCTTACTTGGGATTC 4636
QY 175 CCTGTGAGGATATGAACTCCCACTGAGATGGGTATATCTTCTGTTTAAACAGGATT 234
Db 4637 CTTAGTGAAGATACCTAGTTGAGACAGAAATGATATATTTCTGTGCTTAACCGAAT 4696
QY 235. CCTCGAGGCTAGTGCACCTAAGAGACAGGTTCCAGGCTGTGGTGTATCTGAGCAT 294
Db 4697 CCTCATGGGAGGAAGAACCATTTCTCAAAAGGTCCCAAAACAGTTGTCTTCTGCAACAT 4756
QY 295 GGCCTAGTTGGAGGTGTAGCAACTGGAATTTCCAACTGCGCCCAACATAGCTGGGCTTC 354
Db 4757 GGCCTGTGCGCAGATTTCTAGTAACTGGGTCAAAACCTTGCCCAACAGAGCCTGGGCTTC 4816
QY 355 ATCTGGCAGATGCTGGTGTTCAGCTGTGGATGGGGAACAGCAGGGGAAAACGCTGGTCT 414
Db 4817 ATCTGTGCTGCTGGTGTTCAGCTGTGGATGGGGAACAGCAGAGGAAAATACCTGGTCT 4876
QY 415 CGAAACACAGACACTCTCCATAGACCAAGATGATGCTGGGCTTCAGTTATGATGAG 474
Db 4877 CGAAACATAGACACTCTCAGTTTCTCAGGATGAATTTCTGGGCTTTCAGTTATGATGAG 4936
QY 475 ATGGCTAGGTTTGGACCTCTCAGTGTATAACTTTTATTTGCAAAAACGGCCAGAA 534
Db 4937 ATGGCAAAATATGACCTACAGCTTCCATTAACCTTCTGAAATAAACTGGCCAAAGAA 4996
QY 535 AAGATCTATTATGTCGGCTATTACAGGGCACCACTGGGCTTTTATGCAATTTTCCACC 594
Db 4997 CAAAGTGATTAATGTCGGCTATTCTCAAGGCCACCACTAGGTTTATAGCAATTTTCACAG 5056
QY 595 ATGCCAGAGCTGGCTCAGAAAATCAAAATGATTTTGTCTTAGACCCATAGCACTGTT 654
Db 5057 ATCCCTGAGCTGGCTAAAAGGATTAATGTTTGTGCTGGGCTGCTGGTGGCTTCCGTC 5116
QY 655 AAGCATGCAAAAAGCCCGGACCAAAATTTTGTGCTGCCAGATATGATGATCAAGGGA 714
Db 5117 GCCTTCTGCTAGCCCTATGGCCAAATTTAGACCATTTACAGATCATCTCATTAAGAC 5176
QY 715 TTGTTTGGCAAAAAGATTTCTGTATCAGACCAAGATTTCTCAGACAACTTTGTTATTAC 774
Db 5177 TTATTTGGAGACAAAGAAATTTCTCCCAAGAGTGGGTTTTTGAAGTGGCTGGGTACCCAC 5236
QY 775 CTTTGTGGCCAGGTGATTTCTGATCAGATTTCTAGTATATATCATGTTACTCTGGTGGGA 834
Db 5237 GTTTGCACTCATGCTACTGAGGAGCTCTGTGGAATCTCTGTTTCTCTGTGTGGA 5296
QY 835 TTCAACACCAACAAATATGAACATGAGCCGAGCAAGTGTATATGTCGCCACACTCTTGT 894
Db 5297 TTTAATGAGAGAAATTTAAATATGCTAGAGTGGATGTATATACAAACATTTCTCTGCT 5356
QY 895 GGAACATCTGTGCAAAATTTCTCACTGGAGCCAGGCAAGTGAATCTGGTGAATCTCCG 954
Db 5357 GGAACCTTCTGTGCAAAACATGTTTACCTGGAGCCAGGCTGTTTAAATTTCAAAAGTTTCAA 5416
QY 955 GCATTTGACTGGGGAGTGAGACCAAAATCTGGAATAATGCAATCAGCAACTCTCTGTA 1014
Db 5417 GCCTTTGACTGGGAGGAGCTGCTGCAAGATTTATTTTCAATACACAGAGTTATCTCCC 5476
QY 1015 AGGTACAGAGTCAGAGATATGACGGTCCCTCAGCAATGTGGACAGGAGGTGAGACTGG 1074
Db 5477 ACATACAAATGTGAAGACATGCTTGTGCGGACTGCACTGTGAGCGGGGTGACGACTGG 5536

QY 1075 CTTTCATCCAGAGAGCGTGAATGCTCTCTCTGAGTGACCAACCTCTCTACCAT 1134
 Db |||||
 QY 5537 CTTGAGATGTTACGAGCGTCAATATCTTACTGACTCAGATCAACCACTTGGTGTTCCAT 5596
 Db |||||
 QY 1135 AAGAATATTCCTGAATGGCTCAGTGGATTTTCATCTGGGTTTGGATGCTCTCACCGT 1194
 Db |||||
 QY 5597 GAGAGCATTCGGGAATGGGAGCATCTTGAATCTTCACTTGGGCTGGATGCCCTTGGAGG 5656
 Db |||||
 QY 1195 ATGTACATGAATCATCCATCTGATGCAGCAGAGGAGCAACCT 1241
 Db |||||
 QY 5657 CTTTATAATAAATATTATCTAATGAGGAATATCAGTGAAGCT 5703
 Db |||||

RESULT 11
 US-09-971-392-116
 ; Sequence 116, Application US/09971392
 ; Publication No. US20030134283A1
 ; GENERAL INFORMATION:
 ; APPLICANT: Peterson, David P.
 ; APPLICANT: Pearson, Benjamin G.
 ; APPLICANT: Cocks, Benjamin G.
 ; TITLE OF INVENTION: GENES REGULATED IN DENDRITIC CELL DIFFERENTIATION
 ; FILE REFERENCE: PA-0029 US
 ; CURRENT APPLICATION NUMBER: US/09/971.392
 ; CURRENT FILING DATE: 2001-10-03
 ; PRIOR APPLICATION NUMBER: 60/237,652
 ; PRIOR FILING DATE: 2000-10-03
 ; NUMBER OF SEQ ID NOS: 260
 ; SOFTWARE: PERL Program
 ; SEQ ID NO 116
 ; LENGTH: 2853
 ; TYPE: DNA
 ; ORGANISM: Homo sapiens
 ; FEATURE:
 ; NAME/KEY: misc feature
 ; OTHER INFORMATION: Template ID: 410917.7
 ; NAME/KEY: unsure
 ; LOCATION: 207-395
 ; OTHER INFORMATION: a, t, c, g, or other
 US-09-971-392-116

Query Match 41.4%; Score 525.4; DB 13; Length 2853;
 Best Local Similarity 66.6%; Pred. No. 3.3e-161;
 Matches 751; Conservative 0; Mismatches 376; Indels 0; Gaps 0;
 QY 115 AAAGCTGTGGACCCAGAGCAATTCATGAATATTAGTGAATCATCCAAATCATCAAGGCTAT 174
 Db |||||
 QY 475 ACAGCTGTGGATCTCTGAAACAAACATCATGATGTGATGAATTAATCTTACTGGGATTC 534
 Db |||||
 QY 175 CCCTGTGAGGATATGAGTGGCAACTGAAGATGGGTATATCTTCTGTTAACAGATTT 234
 Db |||||
 QY 535 CCTAGTGAGGAATACCTAGTTGAGACAGAGATGATATATCTGTGCCCTTAACCGAAT 594
 Db |||||
 QY 235 CCTGAGGCTTAGTGCACCTTAAGACAGAGTTCCAGGCTGTGGTGTCTTCTGACGAT 294
 Db |||||
 QY 595 CCTCATGGAGGAGAGACCATCTCTGACAAAGTCCCAACAGTTGTTCTTCTGCAAT 654
 Db |||||
 QY 295 GGCTAGTTGAGGTGTAGCAATGATTTCCAACTGCTGCCCAACATAGCCCTGGGCTTC 354
 Db |||||
 QY 655 GGCTTGTGTCAGATTTCTAGTAACTGGGTACAAACCTTCCCAACAGAGCTGGGCTTC 714
 Db |||||
 QY 355 ATTCGTGACATGCTGTTTGTGAGTGGGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 414
 Db |||||
 QY 715 ATTCGTGATGCTGTTTGTGAGTGGGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 774
 Db |||||
 QY 415 CGAAAAACACAGACACTCTCAGTTCTCAGGATGAATTTCTGGGCTTTCAGTTATGATGAG 474
 Db |||||
 QY 775 CGAAAAACATAGACACTCTCAGTTCTCAGGATGAATTTCTGGGCTTTCAGTTATGATGAG 834
 Db |||||
 QY 475 ATGCTAGTTTGAACCTTCTGAGTGAATTAATTTTTCAGAAAAACGGGCGGAGAA 534
 Db |||||
 QY 835 ATGCAAAATATGACCTTACCAGCTTCCATTAACCTTCCATTCGATTAATAAACTGGCCAGAA 894
 Db |||||

QY 535 AAGATCTATTATGCTGGCTATTTCACAGGCGACACCATGGCTTTATTGCACTTTCCACC 594
 Db |||||
 QY 895 CAAGTGTATTATGTTGGGTCTATTCTCAAGGCGACCACTATAGTTTATTAGCATTTTCACAG 954
 Db |||||
 QY 595 ATGCCAGAGCTGGCTCAGAAAAATCAAAATGATTTTCTTTAGCACCCTAGCAGCTGTT 654
 Db |||||
 QY 955 ATCCCTGAGCTGGCTAAAAGGATTTAAATGTTTTCCTGGCTCTGTTGGCTTCCGTC 1014
 Db |||||
 QY 655 AAGCATGCAAAAGCCCGGAGCAAAATTTTGTGCTGCAGATATGATGATCAAGGA 714
 Db |||||
 QY 1015 GCCTTCTGTACTAGCCCTATGCCCCAATTAGAGATTAACAGATCATCTCAITTAAGGAC 1074
 Db |||||
 QY 715 TTGTTTGGCAAAAGAAATTTCTGTATCAGACCAAGATTTCTCAGACCAACTTTGTTATTAC 774
 Db |||||
 QY 1075 TTATTTGGAGACAAGAAATTTCTCCCGAGAGTGGCTTTTGAAGTGGCTGGGTACCCAC 1134
 Db |||||
 QY 775 CTTTGTGCCAGGTGATTTCTGTATCAGATTTTGTAGTAAATATCATGTTACTTCTGGTGGGA 834
 Db |||||
 QY 1135 GTTTGCACTCATGTACATCTGAAGGAGCTCTGTGGAAATCTCTGTTTCTTCTGTGTGGA 1194
 Db |||||
 QY 835 TTCAACCAACAATATGACATGAGCGGAGCAAGTGTATATGCTGCCACACACTCTTGTCT 894
 Db |||||
 QY 1195 TTTAATGAGAGAAATTTAAATATGTTCTAGAGTGGATGTATATACACACATTTCTCTGCT 1254
 Db |||||
 QY 895 GGAACATCTGTGCAAAATATTCTACACTGGAGCGAGCAAGTGAATTTCTGTGAACCTCGG 954
 Db |||||
 QY 1255 GGAACCTTCTGTGCAAAACATGTTACACTGGAGCGAGCTGTAAATTTCCAAAAGTTCAA 1314
 Db |||||
 QY 955 GCATTTGACTGGGGAGTGGAGCAAAATCTGAAAAATGCAATCAGCCCACTCTCTGTA 1014
 Db |||||
 QY 1315 GCCTTTGACTGGGGAAGCAGTGCACCAAGATTTATTTTATACACCAAGAGTTTCTCTCCC 1374
 Db |||||
 QY 1015 AGGTACAGAGTCAGAGATATGACCGTCCCTACAGCAATTTGGGAAGGAGTTCAGGACTGG 1074
 Db |||||
 QY 1375 ACATACATGTGAAGGACATCTTGTGCGACTGCTGAGCGGGGTTCAGGACTGG 1434
 Db |||||
 QY 1075 CTTTCAATCCAGAGAGAGCTGAAAAATGCTGCTCTCTGAGGTGACCAACCTCATCTACAT 1134
 Db |||||
 QY 1435 CTTGAGATGCTACGAGCTCAATATCTTACTGACTCAGATCAGATCACCACCTTGGTGTCCAT 1494
 Db |||||
 QY 1135 AAGAAATTTCTGAATGGGCTCAGTGGATTTCTATCTGGGTTTGGATGCTCTCTCACCGT 1194
 Db |||||
 QY 1495 GAGAGCATTCGGAATGGGAGCATCTTGACTTCAATTTGGGGCTGGATGCCCTTGGAGG 1554
 Db |||||
 QY 1195 ATGTACAATGAATCATCTCATCTGATGAGGAGGAGGAGCAACCT 1241
 Db |||||
 QY 1555 CTTTATAATAAATTTATTAATCTAATGAGGAATATCAGTGAAGCT 1601
 Db |||||

RESULT 12
 US-10-108-260A-1262
 ; Sequence 1262, Application US/10108260A
 ; Publication No. US20040005560A1
 ; GENERAL INFORMATION:
 ; APPLICANT: HELIX RESEARCH INSTITUTE
 ; TITLE OF INVENTION: No. US20040005560A1el full length cdna
 ; FILE REFERENCE: H1-A0106
 ; CURRENT APPLICATION NUMBER: US/10/108,260A
 ; CURRENT FILING DATE: 2002-03-27
 ; NUMBER OF SEQ ID NOS: 5458
 ; SOFTWARE: Patent In Ver. 2.1
 ; SEQ ID NO 1262
 ; LENGTH: 2506
 ; TYPE: DNA
 ; ORGANISM: Homo sapiens
 US-10-108-260A-1262

Query Match 36.4%; Score 462.2; DB 12; Length 2506;
 Best Local Similarity 67.0%; Pred. No. 1.9e-140;
 Matches 656; Conservative 0; Mismatches 323; Indels 0; Gaps 0;
 QY 263 CAGGTTCAGGCTGTGCTGTGTTACTGACGATGCCCTAGTTGGAGGTGCTAGCAACTGGA 322
 Db |||||

Db 280 CTGGTCCCAACACAGTTGCTTCTGCTCAACATGGCTCGCTGCGCAGATCTTAGTAACCTGGG 339
Qy 323 TTTCACACTGCCCAACATAGCTGGGCTTCACTTCTGCGCAGATGCTGGTTTGAAGTGT 382
Db 340 TCACAAACCTTGCACACAGCCTGGGCTTCACTTCTGCTGATGCTGGTTTGAAGTGT 399
Qy 383 GGATGGGGAACAGCAGCGGGAACCGCTGCTCTCGAAGAACACAGACACTCTCCATAGACC 442
Db 400 GGATGGGGAACAGCAGCGGGAACCTGCTCTCGGGAACATAGACACTCTCAGTTTCTC 459
Qy 443 AGATGAGTCTGGGCTTTCAGTTATGATGAGATGCTAGGTTTGAAGTGTCTGCGAGTGA 502
Db 460 AGATGAGTCTGGGCTTTCAGTTATGATGAGATGCTAGGTTTGAAGTGTCTGCGAGTGA 519
Qy 503 TAAACTTTATTTGTCAGAAACCGGCGCAGGAAAGATCTATTATGTCGGCTATTCCACAGG 562
Db 520 TTAACCTTCACTTGATTAACACTGGCCAGACAGAGTGTATTATGTTGGTCACTTCTCAG 579
Qy 563 GCACCAACCTAGGCTTTATGCAATTTCCACCATGCCAGAGCTGGCTCAGAAATCAAAA 622
Db 580 GCACCAACCTAGGCTTTATAGCATTTTACAGATCCCTGAGCTGGCTAAAGGATTAAAA 639
Qy 623 TGTATTTCTCTTAGCACCATAGCCTGTTAAGCATGCAAAAAGCCCCGGGACCAAT 682
Db 640 TGTATTTCTCTTAGCACCATAGCCTGTTAAGCATGCAAAAAGCCCCGGGACCAAT 699
Qy 683 TTTGTGCTGCGCAGATATGATGATCAAGGGATTTGTTGGCAAAAAGAAATTTCTGTATC 742
Db 700 TAGGACGATTTACAGATCATCTCATTAAGGACTTATTTGGAGACAAAGAAATTTCTTCCC 759
Qy 743 AGACAGATTTCTCAGACAACTGTTATTTACCTTTTGGCGCAGAGTGAATTTCTGTATCAG 802
Db 760 AGAGTGCCTTTTGAAGTGGCTGGGTACCCAGCTTTGCACTCATGCTACTGTAAGGAGC 819
Qy 803 TTTGTAGTATATCATGTTACTTCTGGGTGGATTTCAACACCAACATATGAACATGAGCC 862
Db 820 TCTGTGGAATCTCTGTTTCTTCTGTGTGGATTTAATGAGAGAAATTTAATATGTCTA 879
Qy 863 GAGCAAGTGTATGCTGCCCACTCTTTGCTGGCAACATCTGTGCAAAATATTTACACT 922
Db 880 GAGTGTATGTATATACACACATTTCTCTGCTGGAATTTCTGTGCAAAACATGTTTACACT 939
Qy 923 GGAGCCAGCAGTGAATCTGTGACTCGGGGCTTTGACTGGGGGAGTGAGACCAAA 982
Db 940 GGAGCCAGCAGTGTAAATTTCAAAAGTTTCAAGGCTTTGACTGGGGAAGCAGTGCCAGA 999
Qy 983 ATCTGGAATAATCAATCAGCCAACTCTCTGTAAGGTACAGAGTCAAGATATGACGGTCC 1042
Db 1000 ATTATTTTCAATACACAGAGTATCTCTCCACGTACAAATGTGAAGGACATGCTTGTGC 1059
Qy 1043 CTACAGCAATGTGACAGAGGTGAGGACTGGCTTTCAATCCAGAGACGTGAAATGC 1102
Db 1060 CGACTGCACTGTGGAGCGGGGTCAGACTGGCTTGCAGATGCTTACGACGCTCAATATCT 1119
Qy 1103 TGTCTCTCAGGTGACCAACCTCATCTACATAAGAAATTTCTGAAATGGGCTCAGCTGG 1162
Db 1120 TACTGACTCAGATCAACCACTTGGTGTCTCATGAGAGCAATCCGAAATGGGAGCATCTTG 1179
Qy 1163 ATTTCATCTGGGTTTGGATGCTCTCCCGTATGATCAATGAATTCATCTGATGC 1222
Db 1180 ACTTCATTTGGGCTTGGATGCTCTCCCGTATGATCAATGAATTCATCTGATGC 1239
Qy 1223 AGCAGGAGGAGACCAACCT 1241
Db 1240 GGAATATCAGTGAAGCT 1258

RESULT 13

US-10-420-564-3

; Sequence 3, Application US/10420564

; Publication No. US20040001619A1

; GENERAL INFORMATION:

; APPLICANT: Bolen, Paul L.

; APPLICANT: Cihak, Paul, L.
; APPLICANT: Scharpf Jr., Lewis G.
; TITLE OF INVENTION: Recombinant Kid Pregastric Esterase and Methods for
; TITLE OF INVENTION: Its
; TITLE OF INVENTION: Production and Use
; FILE REFERENCE: IFP-0009
; CURRENT APPLICATION NUMBER: US/10/420,564
; CURRENT FILING DATE: 2003-04-22
; PRIOR APPLICATION NUMBER: US/10/043,665B
; PRIOR FILING DATE: 2002-09-24
; PRIOR APPLICATION NUMBER: US 09/186,489
; PRIOR FILING DATE: 1998-11-05
; NUMBER OF SEQ ID NOS: 7
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 3
; LENGTH: 1411
; TYPE: DNA
; ORGANISM: Kid (Goat)
; US-10-420-564-3

Query Match 36.0%; Score 456.8; DB 12; Length 1411;
Best Local Similarity 62.0%; Pred. No. 7.9e-139;
Matches 741; Conservative 0; Mismatches 452; Indels 3; Gaps 1;
Qy 42 AATGGAATGTGGCTTCTGATTTCTGGTGGCTATATGTTCCAGAGAAATGTGAATTCAGT 101
Db 61 ATTGMAATGTGGCTTCTGATTTCTGGTGGCTATATGTTCCATCCACATGCTGGAAATGCATT 120
Qy 102 ACATATGC---CAACTAAAGCTGTGACCCAGAGCAATTCATGAATATTAGTGAATCAT 158
Db 121 TTGTTTCTTGGAAAAATTTGCTAAGAACCTTGAAGCCAGTATGAATGTGAGTCAGATGAT 180
Qy 159 CCAACATCAAGGCTATCCCTGTGAGGAATATGAAGTCGCACTCGAACTGGAAGTGGGTATATCCT 218
Db 181 TTCTTCTGGGGCTACCCAGTGAAGTGCATATAAGTTATACTGCAAGTGGCTATATCCT 240
Qy 219 TTCTGTTAA CAGGATTCCTGAGGCTTAGTGCACCTAAAGAGACAGGTTCCAGGCTGT 278
Db 241 TCAGGTCTATCGGATTCCTCATGGAAGAAATGATGCTAATCATTTAGGTGAGAGACCTGT 300
Qy 279 GGTGTTACTCAGCATGGCTAGTGTGAGGTGCTGCAACTGATGTTCCAACTGCTCCAA 338
Db 301 TGTGTTCTGACGATGGTCTTCTGCTCAGCTCAAACTGGAATTTCAACCTTCCAA 360
Qy 339 CAATAGCTGGGCTTCATTTCTGCGAGATGCTGGTTTGAAGTGGTGGAGAACAGCAG 398
Db 361 CAACAGCTGGGCTTCTCTCTGCGAGATGCTGGTTTATGAGCTGTGCTGGGGAACAGCAG 420
Qy 399 GGGAAAGCCTGTGCTCGAAGAACACAGACACTCTCCATAGACCAAGATGAGTCTGGGC 458
Db 421 AGGAAACACTTGGGCCCAGGAACATTTATFATATTCACAGACTCCCTCGAATTTCTGGGC 480
Qy 459 TTTCAAGTTATGATGAGTGGCTAGGTTTGAACCTTCTGCGAGTGAATAAATTTTATTTTGA 518
Db 481 TTTCAAGTTTGAATGAATGGTGAATGATGACCTTCCATCTACAAATGATTTTATCTTAA 540
Qy 519 GAAAAAGGCGCCAGGAAGAAATCTATTATGTCGCTATTCAGAGGCAACCAATGGGCTT 578
Db 541 GAGAACAGGACAGAGAAAGCTACACTATGTTGGCCATTCCTCCAGGCAACCAATGGGCTT 600
Qy 579 TATTGCAATTTTCCACCATGCGAGCTGCTCGAAGAAATCAAAATCTATTGCTTTTAC 638
Db 601 TGTGCGCTTTTCTACCAATCCCACTGGCTGAAAAATGGAAGTCTTCCATGCAATAGC 660
Qy 639 ACCATAGCCACTGTTTAAAGCATGCAAAAAAGCCCCGGGACCAAAATTTTGTGTGCGCAGA 698
Db 661 CCCAGTCCGACAGTGAAGCAGACACCCAGAGCCTGTTTAAACAACTTGCCTATTCTCTCA 720
Qy 699 TATGATGATCAAGGATTTGTTGGCAAAAGAAATTTCTGTATCAGACAGATTTTCTCAG 758
Db 721 CTTCTCTCTTCAAGATTTATTTGGTTAAACAAATGTTTCTACCCACACAAATTTTGTGAACA 780
Qy 759 ACAACTGTTATTTAGCTTTGTTGGCCAGGCTGATTTCTTGATCAGATTTGTAGTAATCAT 818


```

RESULT 15
US-10-420-564-4
; Sequence 4, Application US/10420564
; Publication No. US20040001819A1
; GENERAL INFORMATION:
; APPLICANT: Bolen, Paul L.
; APPLICANT: Cihak, Paul L.
; APPLICANT: Scharpf Jr., Lewis G.
; TITLE OF INVENTION: Recombinant Kid Pregastric Esterase and Methods for
; TITLE OF INVENTION: Its
; FILE REFERENCE: IFF-0009
; CURRENT APPLICATION NUMBER: US/10/420,564
; CURRENT FILING DATE: 2003-04-22
; PRIOR APPLICATION NUMBER: US/10/043,665B
; PRIOR FILING DATE: 2002-09-24
; PRIOR APPLICATION NUMBER: US 09/186,489
; PRIOR FILING DATE: 1998-11-05
; NUMBER OF SEQ ID NOS: 7
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 4
; LENGTH: 1134
; TYPE: DNA
; ORGANISM: bovine
US-10-420-564-4

Query Match          35.7%; Score 453.6; DB 12; Length 1134;
Best Local Similarity 62.7%; Pred. No. 7.6e-138;
Matches 705; Conservative 0; Mismatches 419; Indels 0; Gaps 0;

QY 111 AACTAAAGCTGTGGACCCAGAGCAATTCATGAATATTAGTGAATCATCCAAATCAAGG 170
DB 9 AAAAAATTGCTAAGAACCCCTGAAGCCAGTATGAATGTTAGTCAGATGATTTCTACTGGGG 68

QY 171 CTATCCTCTGTGGATATGAGTCGCACTGAGTATGATGATGATGATGATGATGATGATGAT 230
DB 69 CTACCCAGTGTAGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 128

QY 231 GATTCCTCGAGGCTAGTGTCAACCTTAAAGAACAGAGTTCACAGGCTGTGGTGTACTGCA 290
DB 129 GATTCCTCATGAAGAACATATGCTATCAATTTAGGTGAGACCTGTGTGTCTGCA 188

QY 291 GATGCGCTAGTGTGGAGGTGTAGCACTGATGATGATGATGATGATGATGATGATGATGATG 350
DB 189 GCATGCTCTCTTGGATCAGCCACAACTGGATTTTCAACCTGCGCCCAAGAAACAGCCTGGG 248

QY 351 CTTTCATCTCGGAGATGCTGTTTGGACGTGTGGATGAGGACAGCAGGGAACGCGCTG 410
DB 249 CTTCTCTCGGAGATGCTGTTTGGACGTGTGGATGAGGACAGCAGGGAACGCGCTG 308

QY 411 GTCTCGAAACACAGACACTCTCCATAGACCAAGATGATGATGATGATGATGATGATGATGATGA 470
DB 309 GGCCAGGAACATTTATATCTATTCACCACTCCCGGAATTTCTGGGCTTTCAGCTTGA 368

QY 471 TGAGATGGCTAGTGTGTGCTTCTGCTGATGATGATGATGATGATGATGATGATGATGATGATG 530
DB 369 TGAATGGCGGAATATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 428

QY 531 GGAAAGATCTATTATGCTGGGTATTTCAGAGGCGACCACTGAGGCTTTATGCAATTTTC 590
DB 429 GAAGAAGCTACACTATGTTGGCCATTTCCAGGACCACTGATGATGATGATGATGATGATGATG 488

QY 591 CACCATGCCAGAGCTGGCTCAGAAATCAAAATGATGATGATGATGATGATGATGATGATGATGATG 650
DB 489 TACCAGTCCCACTTGGCTGAAAAATCAAAAGTCTTCTATGCAATTAGCCCAAGTGGCCAC 548

QY 651 TGTAAAGCATGCAAAAGCCCGGGACCAAAATTTTGTGCTGCGAGATGATGATGATGATGATGATG 710
DB 549 AGTGAAGTACACAGAGAGCTGTTTAAACAACTTGCACTTATTCCTCTCTTCAA 608

QY 711 GGGATGTTTGGCAAAAAAGAAATTTCTGTATCAGACAGATTTCTCAGACAACTTGTAT 770

```

```

DB 609 GATTATATTTGGTGCACAAAATGTTTACCCACACACACTTTTTCGAAACAATTTTCTGGTGT 668
QY 771 TTACCTTTTGGCCAGGTGATTTCTGATCAGATTTGATGATGATGATGATGATGATGATGATG 830
DB 669 TGAATGTGCTCCCGTGAGACACTGGATGCTCTTTGTAAGAATGCTTGTGTTGCCATTAC 728
QY 831 TGGATTCAACACCAACAATATGAACATGAGCCGAGCAAGTGTATATGCTGCTGCCACACTCT 890
DB 729 TGGATTGACAAATTAATAAATCTTCAACATGATGCTCTTATGATGATGATGATGATGATGAT 788
QY 891 TGCTGGAACATCTGTGCAAAATATTTACACTGAGCCAGGCAAGTGAATTTCTGGTGAAT 950
DB 789 AGCAGGAATCTTCTGTTCAAAAACACCTTCCACTGGAGACAGGCTGTTAAGTCTGGAAAT 848
QY 951 CCGGCAATTTGACTGGGGGAGTGAGACCAAAATCTGGAATAATGCAATCAGCCACTCC 1010
DB 849 CCAAGCTTTTGGCTGGGGAGCCCAATATCAAGAACCTTAATGCAATTCATCAGCCCAACC 908
QY 1011 TGTAAGGTACAGAGTACAGAGATGATGAGGTCCTTACAGCAATGTCGACAGGAGTCAAG 1070
DB 909 TCCCATCTACAATTTAAACAGCCATGATGTCCTCAATTTGCAATGATGATGATGATGATGAT 968
QY 1071 CTGGCTTTCAATCCAGACAGCTGAAATGCTGCTCTCTGAGGTGACCAACCTCATCTA 1130
DB 969 CTTGTTGGCTGACCCCTCAGGATGTTGACTTTCTGCTTTCAAAAACCTCTCTAATCTCAT 1028
QY 1131 CCATAAGATATTTCTGATGGGCTCAGCTGGGATTTTCACTCTGGGCTTTGGATGCTCTCA 1190
DB 1029 CCACAGGAATTTCCAAATTTACAAATCAGTGGACTTTTCTGAGCAATGATGATGATGATG 1088
QY 1191 CCGTATGTACAATGAAATCATTCATCTGATGACAGCGAGGAGAGA 1234
DB 1089 AGAAGTTTACAATGAAATTTGTTCTTTGATGGCCGAAGACAAAA 1132

```

Search completed: February 19, 2004, 14:12:32
Job time : 463.878 secs